

Spatial Objectives in Radiation Therapy Treatment Planning

Arttu Voutilainen

School of Science

Thesis submitted for examination for the degree of Master of
Science in Technology.

Espoo 31.5.2016

Thesis supervisor:

Prof. Harri Ehtamo

Thesis advisor:

D.Sc. (Tech.) Ville Pietilä



Aalto University
School of Science

Author: Arttu Voutilainen

Title: Spatial Objectives in Radiation Therapy Treatment Planning

Date: 31.5.2016

Language: English

Number of pages: 7+53

Department Mathematics and Systems Analysis

Professorship: Systems and Operations Research

Supervisor: Prof. Harri Ehtamo

Advisor: D.Sc. (Tech.) Ville Pietilä

Modern intensity modulated radiation therapy allows a precise control over the dose absorbed by the patient's body. This requires a large number of parameters to define a treatment. Thus instead of specifying the parameters the treatment planner specifies the wanted outcome, after which an optimization algorithm aims to create a treatment plan that satisfies the objectives set by the planner. A common objective type is a dose-volume objective that indicates an upper or a lower limit for dose level for a certain volume of a critical organ, tumor or other structure. Objectives based on biological tissue responses can also be used. However, these objectives do not offer control over the exact 3D distribution of the dose.

This work introduces a new objective type, a spatial objective, to be used in radiation therapy treatment planning, and evaluates its usage in a realistic use case. The objective is implemented on top of the Eclipse Treatment Planning System (Varian Medical Systems, Inc.). A spatial objective sets an upper or a lower limit for the dose in all voxels inside a patient's body. The limit is set for each voxel individually, allowing the planner an exact control over the spatial features of the dose distribution.

In this work it is shown that the spatial objective can be used to direct the optimizer to a physically realistic dose distribution. Spatial objectives can also be utilized to perform plan conversion, i.e., to convert a treatment plan from one treatment machine to another, possibly dosimetrically inequivalent machine. This kind of conversion is required for example if the original treatment machine breaks down. A conversion utilizing the spatial objectives is shown to produce plans that are of comparable or superior quality than those produced by the current tool used in Eclipse.

The results of this thesis can be used to improve the radiation therapy treatment planning and the treatments received by the patients.

Keywords: radiation therapy, optimization, objective, gradient search, treatment planning, plan conversion

Tekijä: Arttu Voutilainen

Työn nimi: Paikkakohtaisen annostavoitteen käyttö sädehoidon suunnittelussa

Päivämäärä: 31.5.2016

Kieli: Englanti

Sivumäärä: 7+53

Matematiikan ja systeemianalyysin laitos

Professuuri: Systeemi- ja operaatiotutkimus

Työn valvoja: Prof. Harri Ehtamo

Työn ohjaaja: TkT Ville Pietilä

Moderni intensiteettimoduloitu sädehoito mahdollistaa potilaaseen kohdistuvan säteilyannoksen tarkan hallinnan. Samalla hoidon määrittämiseen tarvittavien parametrien määrä kasvaa varsin suureksi. Yksittäisten parametrien määrittelemisen sijaan hoidon suunnittelussa määritellään haluttu lopputulos, jonka jälkeen optimointialgoritmi pyrkii luomaan hoitosuunnitelman, joka toteuttaa annetut tavoitteet. Tyypillisesti suunnittelussa käytetään annos-tilavuus-tavoitteita, jotka osoittavat ylä- tai alarajan annokselle tietyssä osassa tärkeän elimen, kasvaimen tai muun rakenteen tilavuutta. Kudoksen biologisiin ominaisuuksiin perustuvia tavoitteita on myös mahdollista käyttää. Nämä työkalut eivät kuitenkaan tarjoa mahdollisuutta hallita annosjakauman kolmiulotteista ominaisuuksia.

Tässä työssä esitellään uusi tavoitetyyppi, paikkakohtainen annostavoite, ja arvioidaan sen käyttöä realistisessa käyttötapauksessa. Työtä varten paikkakohtainen annostavoite toteutettiin Eclipse-hoidonsuunnitteluohjelman (Varian Medical Systems, Inc) yhteyteen. Paikkakohtaisessa annostavoitteessa potilaan kehon jokaiseen tilavuusalkioon voidaan määrätä säteilyannos, jota ei tule ylittää tai alittaa. Annos määrätään jokaiseen tilavuusalkioon erikseen, mikä mahdollistaa annosjakauman kolmiulotteisten piirteiden hallinnan.

Työssä näytetään, että paikkakohtaisia annostavoitteita käyttäen optimointialgoritmin voi ohjata kohti haluttua, fysikaalisesti realistista annosjakamaa. Tavoitteita voi käyttää myös yhdelle hoitokoneelle suunnitellun hoidon muuntamiseen toiselle, dosimetrisesti erilaiselle hoitokoneelle sopivaksi. Muuntaminen on tarpeen esimerkiksi alkuperäisen hoitokoneen rikkoutuessa. Paikkakohtaisia annostavoitteita hyödyntävän muunnoksen näytetään tuottavan samanveroisia tai parempia hoitosuunnitelmia Eclipsen nykyiseen muuntotyökaluun verrattuna.

Tämän työn tuloksia hyödyntäen voidaan kehittää sädehoitojen suunnittelua ja parantaa potilaiden saamaa hoitoa.

Avainsanat: sädehoito, optimointi, optimointitavoite, gradienttihaku, hoidon suunnittelu

Preface

First of all, I want to thank Varian Medical Systems Finland for the opportunity to work on this subject. Radiation therapy is an interesting field to study, and the fact that any improvement in treatment planning can help save someone's life provides great motivation. I am most grateful to my advisor Ville Pietilä for giving me good advice and feedback during the whole process. I also want to thank Prof. Harri Ehtamo for supervision.

I want to thank Varian's Jukka Suominen for giving me time to work on the thesis, Esa Kuusela for corrections and improvements on the text and all the co-workers for helping and teaching me during my time at Varian. Special thanks to Janne Sauvala for sharing a room and keeping me company.

Thank you to my friends at Aalto University, at Teekkarispeksi, and elsewhere, for teaching me about life, love and everything. Finally, I want to thank my family for always supporting and encouraging me.

Espoo, 31.5.2016

Arttu Voutilainen

Contents

Abstract	ii
Abstract (in Finnish)	iii
Preface	iv
Contents	v
Symbols and abbreviations	vii
1 Introduction	1
2 Theory and practice of radiation therapy	3
2.1 Biological background of radiation therapy	3
2.2 Structure delineation	4
2.3 Treatment machines	5
2.4 3D Conformal Radiation Therapy	6
2.5 Intensity Modulated Radiotherapy	6
2.6 Volumetric Modulated Arc Therapy	7
3 Criteria for evaluating treatment plans	9
3.1 Isodoses and color washes	9
3.2 Dose-volume histogram, dose at volume, and volume at dose	10
3.3 Equivalent uniform dose	11
3.4 ICRU Recommendations	12
4 Overview of optimization	14
4.1 Dose calculation	14
4.2 Objectives	15
4.2.1 Dose-volume objective	16
4.2.2 Normal Tissue objective	16
4.2.3 gEUD objectives	17
4.3 IMRT optimization	17
4.4 Other optimization methods	18
5 Spatial objectives in treatment planning	19
5.1 Spatial objective definition	19
5.2 Testing with a simple plausible dose on a model patient	20
5.3 Testing spatial objectives with realistic treatment plans	23
5.3.1 Test setup	23
5.3.2 The metrics used to analyze results	25
5.3.3 Results	25
5.3.4 Effect of spatial objectives in iteration time	27
5.3.5 Choosing an optimal margin	28
5.3.6 Choosing an optimal priority	31

6	Plan conversion using spatial objectives	33
6.1	Test set of treatment plans	33
6.2	Test setup	33
6.3	Plan conversion for prostate cases	35
6.4	Plan conversion for head and neck cases	39
6.5	Additional remarks for further research	42
7	Conclusions	43
	References	45
A	Plan Conversion tests: Prostate	48
B	Plan Conversion tests: Head and Neck	50

Symbols and abbreviations

Abbreviations

3DCRT	3D Conformal Radiation Therapy
API	Application Programming Interface
BEV	Beams-eye View
CT	Computer Tomography
DVH	Dose Volume Histogram
ESAPI	Eclipse Scripting API
EUD	Equivalent Uniform Dose
HI	Homogeneity Index
ICRU	International Commission on Radiation Units and Measurements
IMRT	Intensity Modulated Radiation Therapy
MIP	Mixed integer program
MLC	Multileaf collimator
MRI	Magnetic Resonance Imaging
NTCP	Normal Tissue Complication Probability
NTO	Normal Tissue Objective
OAR	Organ at Risk
PTV	Planning Target Volume
RMSE	Root Mean Square Error
RT	Radiation Therapy
TCP	Tumor Control Probability
TPS	Treatment Planning System
VMAT	Volumetric Modulated Arc Therapy

1 Introduction

Cancer is a major problem in today's world. In 2012, there were 14 million cancer incidences and 8 million cancer deaths (Stewart and Wild [2014]) worldwide, with cancer being the second most common cause of death after cardiovascular diseases (17 million deaths in 2013) (Naghavi et al. [2015]). In the United States, it has been estimated that about 40% of the population will be diagnosed with cancer during their lifetime (Howlader et al. [2015]). More than half of the patients will receive radiation therapy as part of their treatment (Halperin et al. [2004]).

The aim of radiation therapy is to destroy or harm the cancer by killing cancer cells while minimizing adverse effects on healthy tissue. Radiation therapy planning can be viewed as multi-objective optimization problem, where the objectives are to maximize the probability of cure and to minimize the probability of adverse effects.

Radiation therapy can be external or internal. In internal radiation therapy the radiation source is injected into the patient, while in external radiation therapy the radiation comes from outside of the patient's body. Often the ionizing radiation beam from an external source is delivered by photons, but also electrons, protons, neutrons and ions are possible (Hoppe et al. [1979], Paganetti and Bortfeld [2006], Orecchia et al. [1998]). For the purpose of this work, only external photon beam treatments are considered, but the same principles should apply to other modalities as well.

An external radiation therapy treatment plan consists of many parameters such as the number of beams, their geometric setup and the intensities of the beams. With the advent of sophisticated beam-shaping devices such as the multi-leaf collimator that allows intensity modulation within each beam, the number of parameters has dramatically increased. An example is the *intensity modulated radiation therapy* (IMRT) where each beam is divided into a grid of beamlets, the intensity of which can be individually chosen. The increased number of parameters has led to a transition to *inverse planning*, where the treatment planner specifies the wanted outcome and a computer program calculates the parameters that come closest to having the desired results.

The inverse planning problem cannot be solved analytically, but numerical optimization methods can be used to find an approximate solution (Bortfeld et al. [1992]). Current optimization methods for inverse planning use mainly *dose-volume objectives* that specify the upper or lower bounds for dose in some volume of a specific organ or other structure. Usually an upper dose objective is used for organs at risk, while both an upper and a lower dose objectives are specified for the tumor.

While dose-volume objectives work well in most cases, sometimes an objective containing spatial information can be useful. The planned dose can be visualized as dose-volume histograms and through 2D projections of the 3D dose distribution. Dose-volume objectives have a 1-to-1 correspondence with histograms, allowing the treatment planner to investigate if the objectives were achieved, and possibly tune

the objectives for better results. In some cases the planner might want to control the dose distribution directly, for example to decide on a trade-off between a dose to the tumor and a dose to a critical organ intersecting with the tumor. There is no direct mapping from dose-volume objectives to the dose distribution, and in order to control the spatial features the planner needs to create artificial control structures used only for the optimization. An objective able to convey spatial information to the optimizer would give the planner more direct control over the dose distribution.

Another potential use case for the spatial objective would be the case where a treatment plan needs to be converted to a different machine (*plan conversion*). Each plan is designed specifically for some treatment machine. If the machine breaks down, the plan has to be converted to another machine to allow treating the patient while the broken machine is being repaired. The secondary machine may not be dosimetrically equivalent to the original machine, in which case the conversion is not trivial.

In the plan conversion case, the source plan reflects the clinical aims of the planner. Dose-volume objectives can be used to replicate the dose-volume histograms in the new plan, but in this case the spatial information is lost. Hence, the new plan may not exhibit the spatial features intended by the planner. For example, a high-dose region might move from inside the tumor closer to some critical organ resulting in a higher possibility for irradiating the organ due to organ movement and imperfections in patient positioning. The use of dose-volume objectives requires the planner to select the important organs that are considered in the conversion. Utilizing a spatial objective could help to automate the process.

The current planning paradigm aims at a homogeneous dose to the tumor. In reality the tumors are not homogeneous, so one might want to use a heterogeneous dose based on the biological properties of the tumor. This dose painting method suggested by Bentzen [2005] would also benefit from an easy way to specify different dose levels to different parts of the tumor.

To tackle these issues, this thesis reviews the idea of using spatial dose information as an optimization objective and investigates how it should be used to improve the optimization results. The use of spatial objectives in the plan conversion use case is also analyzed.

The thesis is structured as follows. In Section 2, the theory and practice of radiation therapy is reviewed. Section 3 discusses plan quality metrics that can be used to assess the quality of the treatment plans. The current optimization methods are described in Section 4. In Section 5 the spatial objective is presented and studied. Section 6 analyzes the plan conversion use case, and Section 7 presents the conclusions.

This thesis is written at Varian Medical Systems Finland, Helsinki. All of the test data is provided by Varian and a modified version of Varian's Eclipse Treatment Planning System is used to perform the tests.

2 Theory and practice of radiation therapy

Normally the lifetime of a cell, including its division and death, is regulated. This means that a cell divides only when new cells are needed and dies through apoptosis when it is no longer needed. If these regulations cease to work, the cell starts to divide in an uncontrolled manner and fails to self-eliminate, which results in a *tumor* being created. (Hall and Giaccia [2006])

If the tumor grows by displacing the healthy tissues next to it, it is called benign. If treatment is needed, such tumor can often be surgically removed. Cancer tumors are always *malignant*, meaning that the tumor grows by invading adjacent tissues (Hall and Giaccia [2006]). Surgical removal of a malignant tumor is complicated and sometimes impossible, as removing the tumor would also necessitate removal of healthy tissue. There might also be microscopic metastasis (cancer cells not directly connected to the primary tumor), the removal of which is difficult or impossible through surgery. The unattended microscopic metastasis can lead to the recurrence of the tumor (Moss and Cox [1994]). For these reasons there is need for a cure which kills cancerous cells but spares the healthy cells.

2.1 Biological background of radiation therapy

Radiation therapy works mainly by using ionizing radiation - photons (X-rays), electrons, protons, neutrons or ions - to damage the cells' DNA. The radiation can damage DNA either directly or indirectly (by creating radical ions that react with the DNA molecules). A single strand break is likely to be repaired by the cell, as well as multiple single strand breaks when they are spatially far apart. In case there are two breaks opposite to each other or in close proximity, the probability of a double strand break is high. Double strand breaks are believed to be the main reason for radiation-induced cell death (Hall and Giaccia [2006]). Healthy cells are better at repairing damage than cancer cells. This combined with other effects (such as reoxygenation of tumor cells and reassortment of cells within the cell cycle) leads to the tumor cells being more sensitive to radiation when the radiation dose is delivered in small fractions over longer period of time instead of as one large fraction at once. Nevertheless, the normal tissue may be unable to repair all damage to the DNA strand or they may make erroneous corrections, leading to complications.

The probability that a certain dose of radiation kills the cancer cells is called the *tumor control probability (TCP)* (Hasson et al. [2013]). Likewise, the probability of complications in healthy tissue is known as the *normal tissue complication probability (NTCP)* (Michalski et al. [2013]). The aim is to have a dose for which the TCP is high while the NTCP stays low. The ratio of TCP and NTCP is known as the therapeutic ratio. The therapeutic ratio, TCP, and NTCP are schematically illustrated in Figure 1.

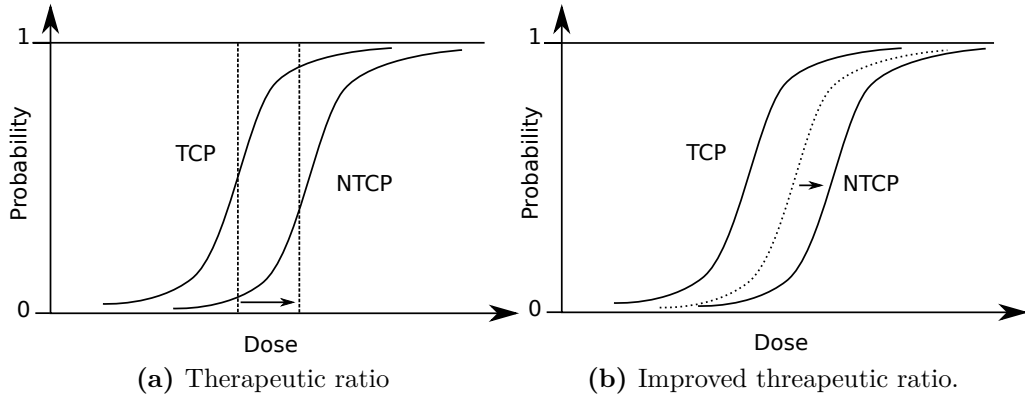


Figure 1 – The therapeutic ratio is a function of the tumor control probability and the normal tissue complication probability. (a) Increasing the dose leads to a higher TCP but also to a higher NTCP. (b) Improvements in the treatment separate the gap between the TCP and the NTCP.

Improving the therapeutic ratio is the main focus of the radiation therapy research. The fractionation pattern (how much dose per fraction, how many fractions and how frequently) affects the therapeutic ratio. Drugs and heating the tumor can be used to increase cancer cells' sensitivity to radiation (Halperin et al. [2004]). The distribution of the dose also has an effect on the therapeutic ratio. Increasing the dose the tumor receives increases the TCP, and reducing the dose that is absorbed by the normal tissues decreases the NTCP.

2.2 Structure delineation

For the purpose of treatment planning it is important to have models of the tumor and healthy tissues. Before a treatment plan (later referred to as *plan*) can be created, the patient must be scanned with a CT or MRI scanner to produce a three dimensional (3D) image of the body. The scanners take 2D-images, called *slices*, that can then be combined to form the 3D representation. The resolution of the images as well as the spacing between the slices control the accuracy of the 3D model. Using the 3D data, the volumes of the tumor and *organs at risk* (OAR) can be defined. These defined volumes are called *structures*.

For treatment planning it is customary to define multiple tumor volumes with the gross tumor volume (GTV) being the known tumor volume, and the clinical target volume (CTV) encompassing the GTV and the subclinical malignant disease. The *planning target volume* (PTV) is a volume created by extending the CTV with a margin corresponding to the uncertainties that come from the treatment machine parameters, tumor movement, and patient setup (Grégoire and Mackie [2011]). There can be multiple tumors leading to multiple PTVs. Multiple PTVs can also be used to implement simultaneous integrated boost strategy, where some parts of the tumor receive a higher amount of dose. The PTV is also referred to as the *target*.

2.3 Treatment machines

Modern radiation therapy treatment machines consist of three main parts: linear accelerator (linac) which produces the radiation, rotating gantry through which the radiation beam is projected and treatment couch on which the patient lies. A modern radiation therapy machine is shown in Figure 2. The intersection of gantry's rotational axis with the beam axes is called the *isocenter*. The couch also has a rotational axis that goes through the isocenter. Using the rotations of the gantry and the couch radiation can be delivered in principle from any solid angle around the isocenter. Often all beams are chosen to be co-planar, in which case only the gantry rotates during treatment.

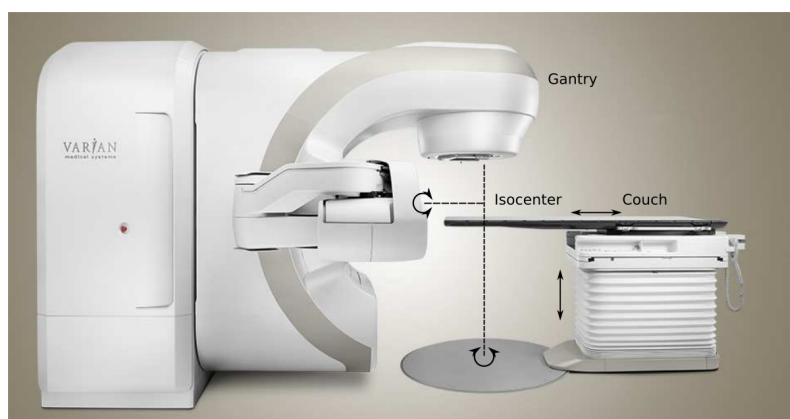


Figure 2 – A modern radiation therapy treatment machine. Image courtesy of Varian Medical Systems, Inc.

To control the shape of the beam, a beam limiting device can be used. Current machines usually include a *multileaf collimator* (see Figure 3) that is able to produce very detailed shapes for the beam (Halperin et al. [2004]). An MLC consists of tens of opposing pairs of leaves that are able to move independently of each other. The leaves are thick enough not to let (almost) any radiation through, limiting the beam to the gaps between the leaves.

The linear accelerators are often capable of producing radiation of different energies (Halperin et al. [2004]). For photon treatments, the energy is given in megavolts (MV). The energy has an effect on the distribution of dose inside the patient. Higher energy radiation penetrates deeper into the tissue. The most common energies for photons are from 6 MV to 18 MV.

For the purpose of treatment planning the beam is usually divided into a grid of *beamlets*. By moving the leaves during treatment it is possible to modulate the *fluence* (radiant exposure) of each beamlet. Beamlets' fluences make a *fluence map* or *intensity map* for the beam. Modern planning systems include a beams-eye view (BEV) that projects patient's 3D image towards the radiation source (Halperin et al. [2004]). Composing the BEV with the fluence map lets the planner see which areas will be included in the beam.

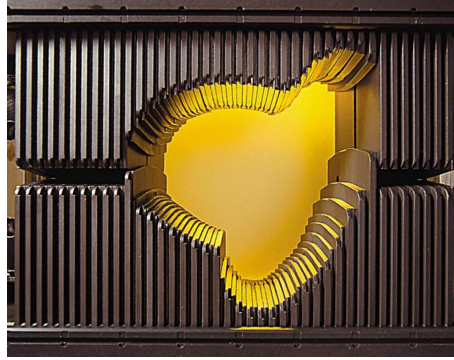


Figure 3 – Varian’s Millennium 80 multileaf collimator can be used to shape the beam. Image courtesy of Varian Medical Systems, Inc.

2.4 3D Conformal Radiation Therapy

In *3D conformal radiation therapy* (3DCRT), the radiation beams are fitted to the projection of the target (Halperin et al. [2004]). After the 3D image of the patient has been acquired, suitable beam directions are chosen by the treatment planner (referred to shortly as *planner*) using the center of the tumor as the isocenter. Some beam limiting device can be used to limit the shape of the beam to the projection of the target, taking into account possible OARs that the planner wants to avoid. 3DCRT planning is an example of a forward-planning method, as it does not require optimization - the planning system calculates the results of the plan, but it is up to the planner to tune the parameters. An example BEV from a 3DCRT plan is shown in Figure 4(a).

2.5 Intensity Modulated Radiotherapy

While the 3DCRT uses uniform fluence for the beams, in *intensity modulated radiotherapy* (IMRT) the fluence of each beamlet is assigned separately. The advantage of IMRT is higher dose conformity, i.e., relatively bigger part of the dose is delivered to the target instead of healthy tissue, allowing the NTCP to be decreased while keeping the TCP high. The non-uniform fluence can be realized with the MLC by moving the leaves during the treatment, thus modulating the time each beamlet is on. Due to limited speed and precision of the MLC, not all fluence maps can be realized exactly.

The higher conformity can be seen especially when the target is concave, e.g., a tumor wrapped around the spinal cord (Grégoire and Mackie [2011]). A downside of IMRT treatments is that the dose homogeneity for the target may be worse than in 3DCRT (see for example Arbea et al. [2010]). Also, due to the accuracy of the beams, the results are more sensitive to patient positioning, patient movement, and organ movement during treatment.

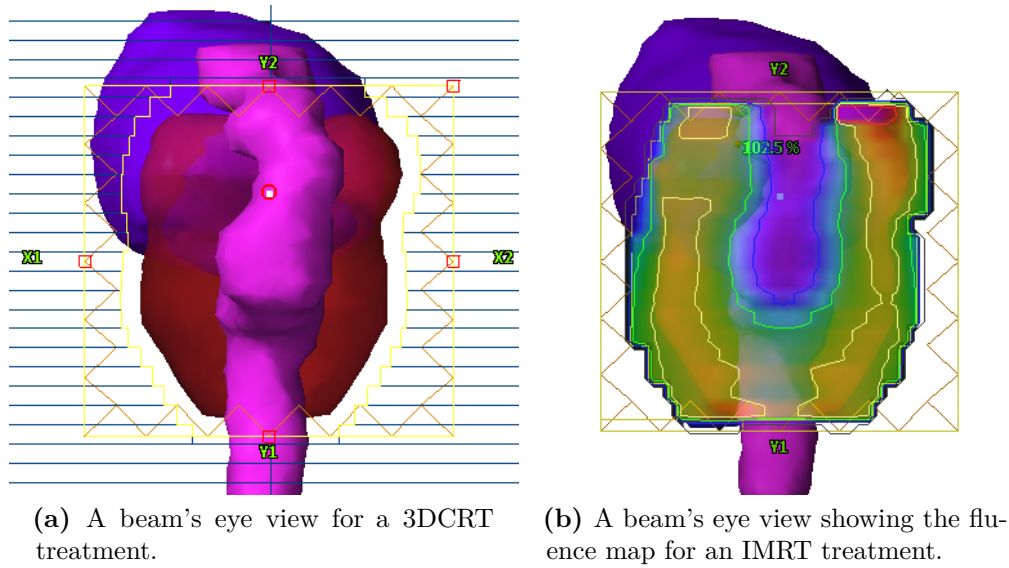


Figure 4 – A comparison of 3DCRT (a) and IMRT (b) BEVs. PTV (red structure) and two OARs (blue and magenta structures) are shown. Using an MLC, the beams used for 3DCRT treatment can be fitted to the projection of the target volume. However, also the OAR (magenta) in front of the PTV receives full dose from this beam. A small margin is used around the PTV to account for imprecision. In IMRT the fluence maps are optimized to produce a dose distribution with certain properties. The fluence map is displayed as a colorwash on the BEV, with red being high fluence and blue low fluence. Higher fluence is assigned to those beamlets that intersect only with the PTV. In reality this kind of beam angle would probably not be chosen for 3DCRT.

Compared to 3DCRT, IMRT treatments have much more parameters that can be chosen or optimized. Hence, IMRT treatments are usually created by inverse planning where the planner defines objectives for the treatment and an optimization algorithm tries to find parameters that result in the wanted outcome. An example fluence map of an optimized IMRT plan is shown in Figure 4(b).

2.6 Volumetric Modulated Arc Therapy

In *volumetric modulated arc therapy* (VMAT), the gantry is rotating while the beam is on and the MLC's leaves are moving. Rotation can be either a full 360° arc, some segment of it, or even multiple arcs. The fluences of the beamlets are modulated by changing the angular velocity of the gantry, changing the intensity of the radiation, and using the MLC. VMAT can be seen as an extension of IMRT with an infinite number of fields but homogeneous fluence per angle. In order to increase modulation per angle, multiple arcs can be used, as the summed up fluence from a given direction consists of the sum of multiple fluence maps. VMAT allows for better OAR sparing in some cases, but may also result in larger volume of irradiated tissue (Teoh et al. [2014]).

VMAT has the advantage of being faster than IMRT. For example, Quan et al. [2012] report typical treatment times of 2.6 min for VMAT plans with two full arcs, compared to 4.7–14 min for 8–24 beam IMRT plans. The difference in the time required to deliver the comparable amount of dose is due to the beam being constantly on in the VMAT treatment as opposed to IMRT where the beam is switched off while the gantry rotates. Shorter treatment time has multiple advantages. Reducing the time it takes to treat one patient allows the machine to treat more patients in a given time, thus reducing the cost of treatment. It is also more comfortable for the patient and there is less patient movement during the treatment which improves the accuracy of the dose delivery.

3 Criteria for evaluating treatment plans

The treatment plan creation is a multi-objective optimization problem, with the high-level objectives being maximal tumor control probability and minimal normal tissue complication probability. As the objectives are conflicting (higher TCP requires more radiation which leads to higher NTCP), the treatment planners must decide on the trade-offs, i.e., which objective is more important and which can be sacrificed to improve the others. (Halperin et al. [2004])

The high-level objectives such as the TCP and the NTCP depend on the biological properties of the tumor and normal tissues. Hence, the calculation of the TCP and the NTCP requires knowledge of biological parameters which are difficult to determine exactly for each patient. Studies have been done to approximate the parameters (for example Emami et al. [1991] and Marks et al. [2010]), but the results so far are inconclusive.

While the use of the TCP and NTCP models in treatment planning is being studied (Bentzen et al. [2010]), it is not yet a standard practice. In order to compare plans, metrics that reduce plans into plain numbers are needed. For that purpose the *International Commission on Radiation Units and Measurements* (ICRU) develops recommendations related to radiotherapy, including quantities that can be calculated from 3D dose distributions and reported.

An important concept relating to the metrics is the *prescribed dose*. The dose prescription tells how much radiation is needed for the treatment to be effective (ICRU [1993]). The prescription can be for example the minimum or mean dose to be absorbed by the PTV. Using the prescribed dose as the norm, other doses can be represented relative to it.

This Section presents methods for visualizing the dose distribution and some commonly agreed criteria that are then used to analyze the results in sections 5 and 6.

3.1 Isodoses and color washes

The treatment planning system calculates the dose on a 3D image of the patient. One way to visualize the dose on a 2D view is to show only a single 2D slice of the 3D image. A color wash is formed by coloring the slice based on the dose in each voxel of the slice.

An *isodose* is a line made of points with equal dose (Halperin et al. [2004]). For example the isodose lines for dose levels 90%, 95%, 100%, 105% and 110% (100% being the prescribed dose) can be drawn to visualize how the dose behaves around the target and if there are any high-dose regions outside the target.

3.2 Dose-volume histogram, dose at volume, and volume at dose

Dose-volume histogram (DVH) is a method to visualize how much dose a given structure receives. The DVH is a 2D graph with volume on the y -axis and dose on the x -axis. The volume is usually shown relative to each structure's volume; dose can be either absolute or relative to the prescribed dose.

A DVH can be drawn as cumulative or differential. A differential DVH shows for each dose the volume that gets exactly that dose, while a cumulative DVH shows for each dose the volume that gets at least that amount of dose (Grégoire and Mackie [2011]). In this work DVH refers to the cumulative DVH. Examples of cumulative and differential DVHs are shown in Figure 5.

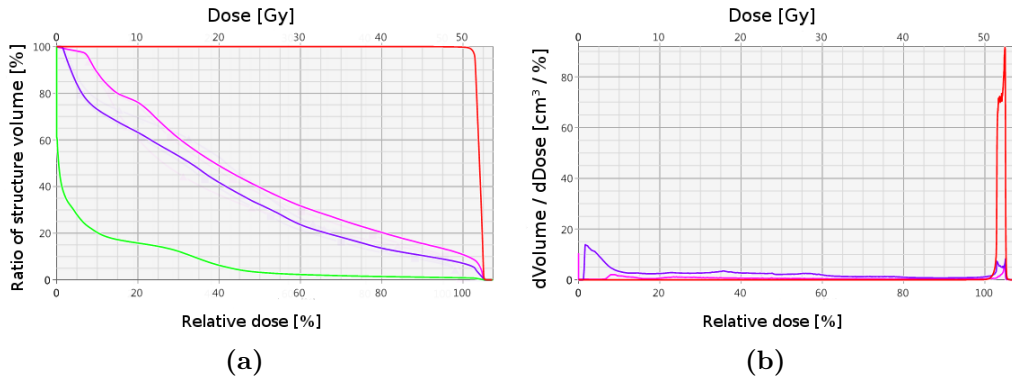


Figure 5 – Examples of cumulative (a) and differential (b) DVHs. The red line is the PTV and the three other lines (pink, purple and green) are organs at risk.

The DVH graph can be used to analyze both the dose for the target and the doses for the organs at risk. For the PTV, the planner usually aims for a homogeneous dose of 100%, which in a DVH corresponds to a line going from (0, 100) to (100, 100) and then dropping to (100, 0). For OARs, the aim is to have the curves decreasing as rapidly as possible. When comparing two DVH curves for an OAR, if one is always below the other, then it is absolutely better. On the other hand, if the two curves intersect, it is not trivial to know which one should be preferred, as it depends on how the organ in question responds to radiation. Examples of DVH comparisons are shown in Figure 6.

The DVH works well for evaluating a single plan, as it gives a good overview of the doses to different structures. Nevertheless, it is devoid of any spatial information. For example, it might make a difference if a high-dose region is deep inside the tumor or close to the border between the tumor and a healthy tissue. In the future there might be a change from a homogeneous PTV dose to a heterogeneous dose based on spatial and biological information about the tumor (Bentzen [2005]), in which case the DVH curve no longer possesses the necessary information to determine the quality of the PTV coverage.

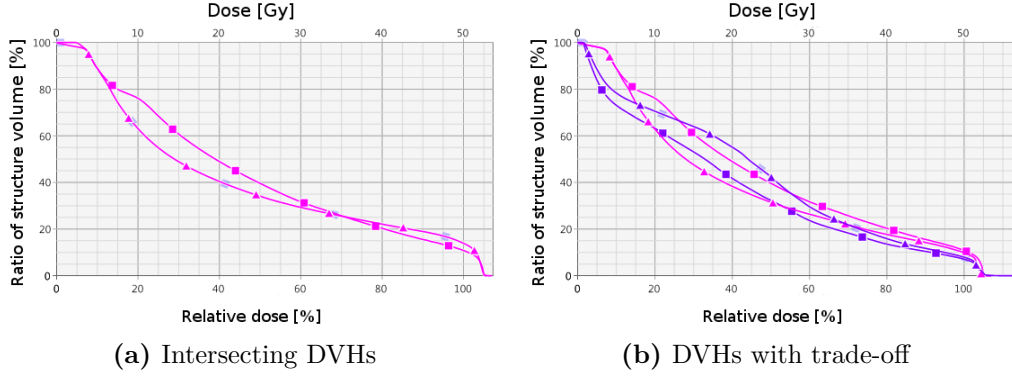


Figure 6 – Examples of DVH comparison between two plans. In (a) the curves for two plans are intersecting, and so it is not possible to say which one would lead to a clinically better result without extra information. In (b) the curves show a tradeoff: the organ drawn in purple receives strictly less dose in plan A (marked with squares) compared to the plan B (marked with triangles), while for the organ drawn in pink the plan B is strictly better than plan A.

Comparing different plans by their DVHs becomes complicated if the plans contain different trade-offs. In such cases the treatment planner typically decides which parts of the DVH are most important for the TCP and NTCP. Automatic comparison of DVHs is difficult for the aforementioned reasons. Moreover, when the number of plans or important structures increases, visual comparison between plans becomes increasingly difficult.

Instead of analyzing at the DVH graph itself, some specific dose-volume values can be calculated. *Dose at volume*, the minimum absorbed dose that covers the specified volume, is usually written as $D_{volume\%} = dose$. Some meaningful values for targets and organs are $D_{0\%}$, $D_{2\%}$, $D_{50\%}$, $D_{98\%}$, $D_{100\%}$ and D_{mean} . $D_{0\%}$ and $D_{100\%}$ represent maximum and minimum doses, respectively, while $D_{2\%}$ and $D_{98\%}$ are called near-maximum and near-minimum.

Volume at dose, usually written as $V_{dose} = volume$, represents the volume that absorbs at least the specified dose. The mean dose to a structure is defined as

$$D_{mean} = \frac{1}{V} \int_0^{D_{max}} D \frac{dV_D}{dD} dD, \quad (1)$$

where $\frac{dV_D}{dD}$ is the differential volume at dose D , i.e., the volume that receives exactly the specified dose.

3.3 Equivalent uniform dose

According to Withers et al. [1988] organs at risk can be classified as *serial* or *parallel*, where serial organs show a threshold response to radiation, while for parallel organs the response is more graded. Examples of serial organs are the spinal cord and the

esophagus. Lung and liver are examples of parallel organs. For a serial organ a dose exceeding a threshold value even in a small volume prevents it from working at all. While radiation can harm a parallel organ, it will not lose all its functionality even if parts of its volume receive a high dose.

Equivalent uniform dose (Niemierko [1997]) and its generalization *generalized equivalent uniform dose* (Niemierko [1999]) attempt to represent the actual dose an organ receives by a dose that, when distributed uniformly to the whole organ, gives the same complication probability. For each organ EUD requires a tissue-specific parameter a , with $a = 1$ giving serial-like results and higher values (for example $a = 100$) being suitable for parallel-like organs. Using $a < 1$ (for example $a = -100$), the formula can also be applied to the target. The problem with EUD is that the results are strongly dependent on the parameter a and it is challenging to determine the exact value of a for each organ. The EUD values are not used in this work.

3.4 ICRU Recommendations

The ICRU report 83 (ICRU83) by Grégoire and Mackie [2011] recommends that for target structures the values $D_{2\%}$ and $D_{98\%}$ to be reported. Near-maximum and near-minimum are suggested instead of absolute maximum and minimum since the real extremes can be highly sensitive to dose calculation resolution and the accuracy of the structure model and delineation. ICRU83 also recommends reporting the $D_{50\%}$ or a similar value, and D_{mean} when deemed clinically relevant - normally for PTV the $D_{50\%}$ ($= D_{median}$) is quite close to the D_{mean} .

In addition to the dose levels, the homogeneity of the dose inside the target is often of interest. For this purpose the ICRU83 suggests the *homogeneity index*:

$$HI = \frac{D_{2\%} - D_{98\%}}{D_{50\%}}. \quad (2)$$

A lower value of HI is preferred, with zero indicating almost perfectly homogeneous dose for the PTV. It is worth noting that the HI information is included in the dose values recommended to be reported. The homogeneity index is only valid when the planner aims for a homogeneous PTV dose.

For critical organs, the relevant dose values depend on whether the organ is serial or parallel. For serial organs the maximum dose ($D_{2\%}$ value) matters the most, while for parallel organs the mean absorbed dose D_{mean} is more relevant. For parallel organs also the volume at a clinically relevant dose can be shown, preferably selected such that going over that dose has high probability of causing serious complications. As no organ is completely serial or parallel, ICRU83 recommends reporting all three values for all organs.

The values calculated for an organ depend on how the organ's volume is delineated. For the purpose of this work, the delineation is not a problem as the comparisons are only between different plans for the same patient with fixed structures. Note that

the organs usually receive a highly inhomogeneous dose, and hence the median dose might differ greatly from the mean dose. The clinically relevant doses for different organs have been studied for example in Emami et al. [1991] and updated in Emami [2013].

Hot spots are defined to be the volumes outside the PTV that receive a dose larger than 100% of the specified PTV dose. (ICRU [1993]). A hot spot with diameter less than 15 mm is in general not considered relevant, unless it is inside a small organ, such as eye or optic nerve. In this thesis only the total volume of the hot spots is calculated and reported as *hot spots outside target* (HOT).

4 Overview of optimization

Traditionally the treatment planning has been a *forward planning* process, where the planner chooses the values for possible parameters (e.g., number of beams and beam directions), calculates the dose, evaluates the plan, and modifies the parameters if necessary. Designing IMRT or VMAT treatments manually is not feasible due to the large number of parameters (e.g., the MLC positions at each time step). For this reason, the modern treatment planning systems are based on *inverse planning* concept: the planner describes the wanted outcome and the planning system calculates the parameters to achieve it.

In IMRT, each beam is discretized into a grid of *beamlets* of fixed size that is usually specified on the isocenter plane. The fluence (intensity) of each beamlet can be set almost independently of other beamlets, giving rise to the term *intensity modulation*. After the fluences for a beam are determined, the corresponding MLC positions can be calculated. Due to the physical limitations of the MLCs, not all fluences are achievable.

According to Bortfeld et al. [1992], dose can be described as

$$d = Dx, \quad (3)$$

where x_k is the fluence of beamlet k , d_i is the resulting dose in voxel i , and D is the dose calculation matrix, with D_{ik} containing the contribution of beamlet k to voxel i . Solving the inverse planning problem would mean inverting the matrix D and calculating $x = D^{-1}d$. However, the matrix is generally not invertible and even if it was, the result could include negative intensities which is clearly not physically plausible. For that end, planning systems include algorithms that can be used to optimize the plan with regards to given objectives. This results in a multi-objective optimization problem and the objectives need to be converted into a single objective function. Typically the user can specify priorities or weights for each objective, and the optimizer sums their weighted costs to produce a single cost value.

The optimization algorithms are typically iterative, improving the solution on every iteration through for example a gradient descent. Other optimization methods are discussed shortly in Section 4.4. The rest of this section is specific to the Eclipse Treatment Planning System and its Photon Optimizer algorithm (Varian Medical Systems, Inc [2015a]) which are used in this work.

4.1 Dose calculation

An important part of the optimization is dose calculation. The treatment plan is usually represented as fluence maps for each beam direction, but all currently used objectives operate on the dose distribution. The dose calculation needs to convert optimized fluences to realistic leaf motion sequences and then to actual doses. In Eclipse, the actual dose is usually calculated using the Anisotropic Analytical

Algorithm (AAA) based on Ulmer and Harder [1995] and developed further in Tillikainen et al. [2008]. The AAA is a Monte Carlo based convolution superposition algorithm, i.e., it uses separate kernels for different kinds of radiation (primary photons, scattered extra-focal photons, scattered electrons) to convolute beamlets and then sums the results to get the final dose.

The whole process of realistic and exact dose calculation is quite slow, and the optimizers usually use a less accurate but faster dose calculation engine. The lightweight dose calculation algorithm used in Eclipse is the Multi-Resolution Dose Calculation (MRDC) algorithm (Varian Medical Systems, Inc [2015a]).

To reduce the difference between the optimizer dose and the final dose, an intermediate dose calculation can be performed to calculate a correction term. The correction term is calculated by subtracting the optimizer dose matrix from the more accurate intermediate dose. The term is then added to the doses from subsequent iterations to correct some of the difference. Using the intermediate dose helps in two ways: it gives the user a more realistic idea of the final dose, and it can make the objectives more meaningful. This is especially important when trying to replicate existing dose distributions. Eclipse allows running the intermediate dose calculation directly during the optimizer, in which case it is based only on the optimized fluences instead of real leaf motions. Other option is to perform accurate leaf motion and dose calculations after optimization and then continue the optimization using the calculated dose as the intermediate dose.

4.2 Objectives

There are three types of objectives: *upper*, *lower* and *exact*. An upper objective means that the specified value should not be exceeded, and is usually used for organs at risk and to restrict the maximum dose inside the target. A lower objective specifies a value that should be achieved. Lower objectives are used to make sure that the target receives a sufficient dose. An exact objective should be matched exactly, and it is used in some cases for targets. Using lower or exact objectives for OARs is in general not suggested as it conflicts with the clinical objective of minimizing the dose to healthy tissues.

The optimization objectives must be differentiated from the clinical goals. While clinical goals aim to specify the outcome of the treatment, optimization objectives are more technical and used to direct the optimizer. However, some clinical goals may be directly convertible to optimization objectives. With the development of more biologically and clinically relevant plan metrics such as the EUD, the TCP, and the NTCP, the role of the clinical goals in the optimization may increase.

Eclipse requires the user to set a priority for each objective. Internally the priorities are converted into weights using a superlinear formula. The weights are then used when the costs and gradients are summed together.

4.2.1 Dose-volume objective

A dose-volume (or DVH) point objective is the most basic type of objective. An upper (lower, exact) point objective indicates that the dose at some volume should not be more (be less, differ) than what is specified in the objective. Using the DVH, the cost of an upper dose-volume point objective is calculated as

$$c(d, v) = \int_{x=v}^{V_d} \frac{1}{2} (D_x - d)^2 dx, \quad (4)$$

where d is the objective dose level, v is the objective volume level, D_x is the actual dose at volume x and V_x is the actual volume at dose x . The formulas for lower and exact objectives are analogous.

In practice the cost is calculated inside the structure for which the objective is specified. For each voxel where the dose is higher (lower) than the specified dose and lower (higher) than D_v , the cost is calculated and added to the total cost of that objective. For the same points, gradients of the cost functions are also calculated and added to the gradient matrix. For a voxel i , the cost and the gradient of an upper point objective are given by

$$c_i(d, v) = \begin{cases} \frac{1}{2} (D_v - D_i)^2, & \text{if } D_v > D_i > d \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

and

$$c'_i(d, v) = \begin{cases} D_v - D_i, & \text{if } D_v > D_i > d \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

respectively, with D_i being the dose at voxel i . The costs and gradients of lower and exact point objectives are similar.

A line objective can be used to replace a continuous group of point objectives associated with the same structure. The advantages of a line objective compared to using many point objectives are that it is faster to calculate and that it does not create unnecessary local optima. The cost and the gradient of a line objective are analogous to the point objective.

4.2.2 Normal Tissue objective

The *normal tissue objective* (NTO) works as a general tool to limit the radiation outside of the target. Its main use is to prevent hot spots, but it can also be used to obtain sharp dose gradients around the target. Eclipse supports manual and automatic NTOs. For a manual NTO, the user has to specify four parameters: margin x_{start} , start dose f_0 , end dose, f_∞ and falloff k . The upper limit for dose at distance x from the border of the target can then be calculated as

$$f(x) = \begin{cases} f_0 e^{-k(x-x_{start})} + f_\infty (1 - e^{-k(x-x_{start})}), & x \geq x_{start} \\ f_0, & x < x_{start}. \end{cases} \quad (7)$$

In case of multiple targets, NTO is calculated for each voxel with regards to the target that yields the highest value. When using the automatic NTO, the software determines the parameters dynamically during optimization with the aim of reducing dose in high dose regions outside the target.

4.2.3 gEUD objectives

The optimizer also supports lower, upper and exact objectives specified through the generalised Equivalent Uniform Dose. Using gEUD in optimization requires the user to specify the biological parameter a used for calculating the gEUD value. The cost arising from an gEUD objective is based on the quadratic difference between the target value and the structure's current gEUD value. The gEUD objectives are not used in this work.

4.3 IMRT optimization

IMRT treatments can be planned using an iterative gradient search optimizer (Varian Medical Systems, Inc [2015a]). The workflow that is iterated is the following:

1. Calculate dose based on fluence maps.
2. Calculate objective costs and gradients based on dose.
3. Project gradients to fluence maps.
4. Optimize fluence map gradient weights using line search.
5. Modify fluence maps accordingly.

During each iteration, the algorithm updates the dose matrix and the gradient matrix. The dose is calculated separately for each field. Field doses are mapped to the dose matrix and summed. Objective costs and gradients are calculated for each voxel related to the objective as specified in Section 4.2. The gradient is projected to the fields' fluence maps and a line search is conducted to find the minimum cost.

After each iteration, the optimizer evaluates a convergence criteria. The criteria can be for example the relative improvement of total cost during the last iteration. If the improvement is smaller than a predefined value, the algorithm is said to have converged.

The optimizer tries to find the optimal fluences for each beam. It does not take into account the physical constraints of the multileaf collimator. The optimizer smooths the fluence maps on each iteration in order to prevent too sharp gradients, thus reducing the difference between the optimal and the actual fluences.

4.4 Other optimization methods

As Langer et al. [2003] note, optimization algorithms can be divided into three categories: those that find some local optima, those that find a local optima with bounds on the error, and those that always find the global optima.

According to Deasy [1997], using dose-volume constraints or optimizing beam angles may lead to the problem having multiple local minima. If only minimum or maximum dose objectives or quadratic deviation from a prescribed dose are used as the cost function, then only a single minimum exists and the gradient search will converge to the global optimum (Webb [1994]). Wu and Mohan [2002] did not find clinically relevant differences between the local and global optima when using dose-volume constraints.

Also other optimization methods are possible. Lee et al. [2003] formulated the IMRT inverse problem, including beam angle optimization, as a *mixed integer program* that can be solved to get a provenly optimal result. In a clinical case they assume to have tens of thousands of constraints and variables, which enforces them to introduce specialized heuristics to make the problem tractable. The heuristic method allows them to produce a feasible solution quite fast, while reaching the provenly optimal solution takes much (20 – 50 times) longer. Still, being able to reach the optimal solution allows to analyze how close the first feasible solution is to the globally optimal solution. In their test cases the difference in objective values ranges from 13.5 % to 33 %, but it is not analyzed if the difference between the feasible and optimal solutions is clinically relevant. No comparison between the MIP and other methods is presented.

Metaheuristic approaches, such as simulated annealing and genetic algorithm, can also be used (Halperin et al. [2004]). These methods explore the solution space in a partly stochastic way, being so able to examine multiple different solutions. The best solution is then chosen, but the methods cannot guarantee that it would be the global optima. The metaheuristics are also much slower than deterministic heuristics such as gradient methods, which discourages their usage.

For VMAT treatments, a different optimization method is required due to the theoretically unlimited number of beam directions and the restrictions for the change in fluencies between consecutive angles. VMAT optimization is not discussed in this work, for more information see for example Otto [2008] and Varian Medical Systems, Inc [2015a].

5 Spatial objectives in treatment planning

A spatial objective is a 3D matrix where each voxel, mapping to a point in the patient’s body, is assigned a dose value. Depending on the type of the objective (upper, lower, or exact), the dose in the corresponding body location should not exceed, be lower than, or differ from the specified value.

Currently none of the major treatment planning systems support spatial objectives as such. The only way to pass spatial information to the optimizer is by creating *optimization structures* (structures with no biological meaning), and assign them dose-volume objectives. For example, the planner may create a structure for a hot spot volume to reduce its dose. More detailed control, such as specifying a gradient, is not feasible using the dose-volume objectives only. In this thesis the spatial objectives are analyzed in the context of Eclipse’s Photon Optimizer algorithm.

The idea of a spatial objective is not new. For example Webb [1994] uses a cost function based on the quadratic difference between a prescribed dose and the current dose. Eclipse’s normal tissue objective is effectively a spatial objective. The difference is in how the spatial objective dose is defined. Webb only uses a single dose level for the target, another dose level for the OARs, and a third dose level for the rest of body. Normal tissue objective’s values are calculated based on the distance to targets. The spatial objective presented here allows the user to set the target levels individually for each voxel.

For the purpose of this work Eclipse was modified to support the spatial objectives through the Eclipse Scripting API (ESAPI). Currently, the API allows the user to query the size, the resolution, and the position of the spatial objective matrix. The user can also set the dose values of the objective matrices, as well as the weights of the objectives. The dose value can also be set to undefined in which case the voxel is skipped when calculating the cost and the gradient of the objective. Only upper and lower spatial objectives are supported, but they can be combined to form an exact spatial objective. The spatial objectives can be used in combination with other objectives.

In this section the aim is to evaluate the spatial objective rather than the accuracy of the optimizer’s dose calculation. Hence, the tests in this section are done using the doses calculated by the optimizer. The effect of the differences in dose calculation is discussed in Section 6.

5.1 Spatial objective definition

During each iteration, the optimizer calculates the cost and the gradient of the spatial objectives by looping through each voxel in the dose matrix and comparing the voxels’ dose values to the corresponding values in the objective matrix.

The cost c_i and the gradient c'_i of the upper spatial objective for voxel i are calculated as follows:

$$c_i(d) = \begin{cases} \frac{1}{2}(D_i - d_i)^2, & \text{if } D_i > d_i \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

$$c'_i(d) = \begin{cases} D_i - d_i, & \text{if } D_i > d_i \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

where D_i is the current dose at voxel i located at (x_i, y_i, z_i) and d_i is the objective dose at the same voxel. The formulas for the costs and gradients of the lower spatial objective are similar:

$$c_i(d) = \begin{cases} \frac{1}{2}(D_i - d_i)^2, & \text{if } d_i > D_i \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

$$c'_i(d) = \begin{cases} D_i - d_i, & \text{if } d_i > D_i \\ 0, & \text{otherwise,} \end{cases} \quad (11)$$

and an exact spatial objective can be created by combining lower and upper spatial objectives with same objective dose:

$$c_i(d) = \frac{1}{2}(D_i - d_i)^2 \quad (12)$$

$$c'_i(d) = D_i - d_i. \quad (13)$$

The total cost of the spatial objective is given by summing the costs of all voxels for which the objective has been set. The gradients are added to the corresponding elements in the optimizer's gradient matrix.

Quadratic cost [Eqs. (8), (10) and (12)] was chosen due to it being used already in the normal tissue constraint as well as the other constraints. The gradient of the quadratic cost is linearly relative to the difference between the targeted and achieved doses [Eqs. (9), (11) and (13)]. A constant gradient resulting from linear cost would not prioritize the voxels based on the dose difference at all, and a superlinear gradient would probably give too much weight to single voxels with large differences.

5.2 Testing with a simple plausible dose on a model patient

The first test is to check if the spatial objectives can direct the optimizer to reproduce a simple but physically plausible dose distribution. Instead of using a real patient data, a cube of uniform tissue was created, with a spherical target structure in its center. This kind of artificial patient is called a *phantom*.

The spherical PTV was given a homogeneous dose through optimization with dose-volume objectives (visualized in Figure 7). The resulting dose was set as the lower and the upper objectives, creating in effect an exact objective.

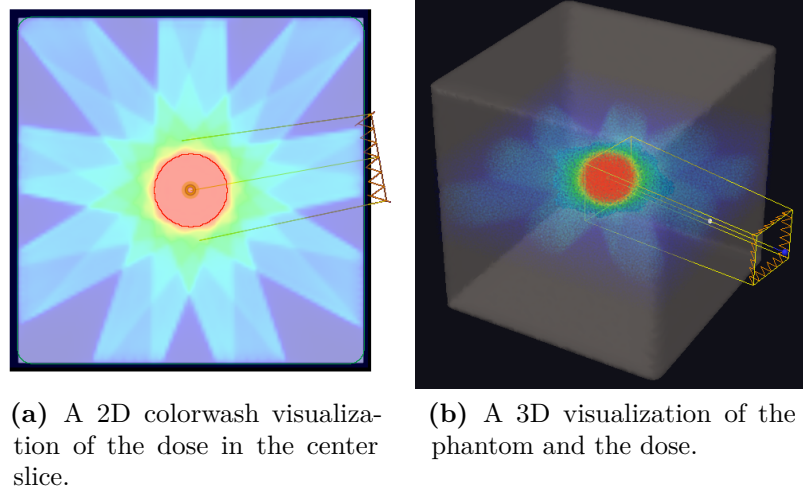


Figure 7 – Figures (a) and (b) visualize the reference dose. The dose is highest in the intersection of the beams (red) and lowest in parts outside of all beams (darker blue). In Figure (a), one of the beams is also drawn for reference. The target is indicated by the red circle.

In the optimizer algorithm the spatial dose objective matrices must coincide with the optimizer dose matrix so that the cost can be calculated for each voxel. If the resolution or position of the spatial objective matrices do not match with those of the optimizer’s dose matrix, the optimizer has to convert the matrix to its internal format by approximating the objective values at missing voxels, using for example trilinear interpolation. As the dose distribution does not change linearly, the interpolated values might differ from the calculated values. If the approximated value is higher than the actual value in a lower objective (or respectively lower in an upper objective), the optimizer will try to increase (decrease) the value at that voxel, but as increasing the dose at one voxel necessarily affects also the neighboring voxels, this causes the optimizer to deviate from the original objective dose.

The first phantom used here had a 5 mm interval between the slices in the dose matrix, while the optimizer uses 2.5 mm slicing, so the optimizer had to interpolate the values for every second slice. In Figure 8(a) it can be seen that in even slices error is negligible, while the odd slices (interpolated) do show differences between the calculated and the interpolated doses. The differences are especially strong in the slices where the dose changes are large and non-linear, for example at the borders of the beams [Figure 8(b)].

Other possibilities for treating missing values would include using one of the neighboring voxels’ value or skipping the missing values entirely. To mitigate the problem outlined above, minimum (maximum) neighbor value from the objective matrix should be used for lower (upper) objective. Skipping non-coinciding voxels could be problematic if the number of such voxels is large, as it results in loss of information. In the worst case, no voxels coincide, leading to the spatial objective being discarded

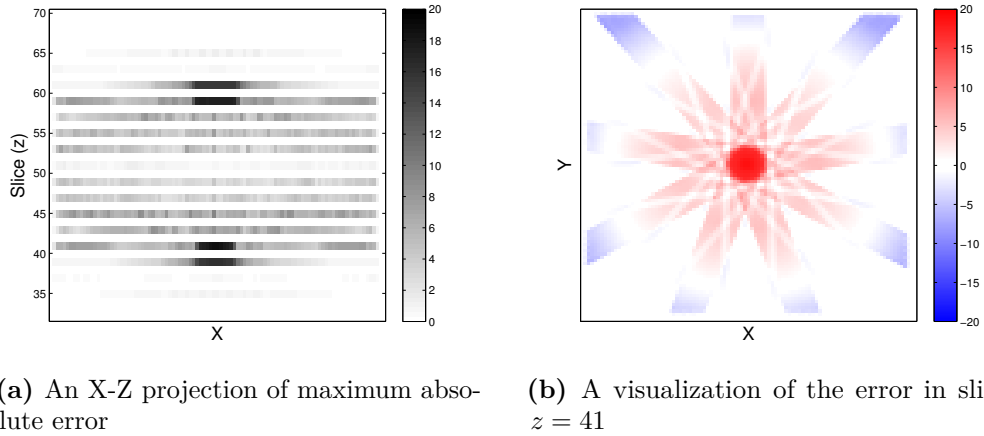


Figure 8 – The difference between interpolated and calculated doses. Panel (a) shows that there are significant differences only on the interpolated (odd) slices. Panel (b) shows the differences on slice 41, on the lower end of the beams, where the differences between interpolated dose and calculated dose are relatively large. The values shown are relative to the prescribed dose.

totally. When possible, the best option is to make sure that the optimizer dose matrix and Eclipse’s objective matrices coincide, in which case there are no missing values. This approach is used for the tests in this thesis.

In this case the problem was fixed by creating another phantom with 2.5 mm slicing interval. A similar dose distribution was created using the same dose-volume objectives, and the resulting dose was set as the spatial objectives.

Using only upper and lower objectives with a dose that is known to be achievable means that at least one zero-cost global minimum must exist in principle. In practice factors such as the fluence smoothing and the fact that the optimizer is an iterative algorithm may prevent the optimizer from reaching the zero-cost minimum. In this case, the optimizer detects convergence after 70 iterations with costs 12.246 and 12.073 for upper and lower spatial objectives, respectively. The reported costs depend on the spatial objective weights and the objective costs and only their relative scale is relevant. These costs are small compared to the initial cost (about 264000 for lower spatial objective), but still significantly larger than zero.

The optimizer’s convergence criteria is triggered if the improvement per iteration is very small. The criteria is a heuristic meaning that significant improvements after the convergence are unlikely, but not impossible. In order to see if the small improvements would sum up to a significant improvement over time, the same plan was optimized up to 1000 iterations.

The logarithm of the total cost after each iteration is shown in Figure 9. Initially the cost decreases very fast. Around iteration 70 the improvements are very small and convergence is detected. Very slow progress continues until around iteration

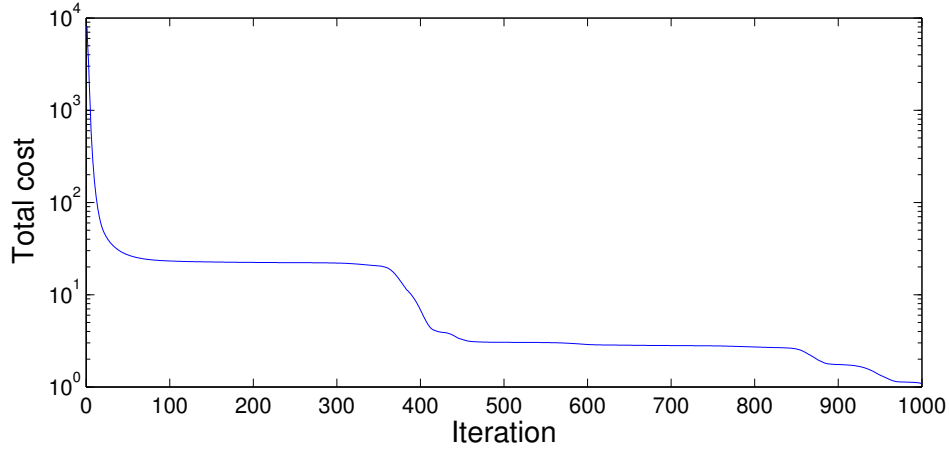


Figure 9 – The total cost of the objective function as a function of the iteration. Note the logarithmic y -axis.

350, where the cost begins to decrease again. This suggests that the algorithm gets stuck in a local optimum or in a zig-zag-motion. After calculating the gradient the optimizer performs a line search in order to find the local minimum. The range for the line search is modified internally at each iteration, and the escapes seemed to be correlated with increased range.

The total cost after 1000 iterations is around 1 and still decreasing on each iteration. These results show that while the optimizer is able to get very close to the exact dose specified by the spatial objectives, it may get stuck in a local optimum or in a zig-zag motion. This suggests that the optimizer algorithm could be improved. For example, allowing the optimizer to occasionally choose a move that increases the cost might allow it to escape the local minima faster. However, improving the optimizer in this respect is outside of the scope of this Thesis. Since the initial result is already very close to the targeted dose, it is unlikely that the local optima would have much effect in the plan conversion case where a zero-cost global optimum does not exist.

These tests were performed using the fluence smoothing. Turning it off seemed to worsen the results. It might be that smoothing the fluence maps smooths also the optimization landscape and allows the gradient descent to find a better solution.

5.3 Testing spatial objectives with realistic treatment plans

5.3.1 Test setup

To test the spatial objectives' capability to guide the optimizer in recreating a realistic dose distribution and to see how different parameters affect the result, two IMRT plans were generated, one for a prostate patient (A) and one for a head and neck patient (B). These are two cancer types commonly treated with external radiation.

Upper and lower point objectives were added for the PTV(s) in both plans. For OARs, upper point and line objectives were added. The manual NTO was also used. The objectives were created using Eclipse’s RapidPlan tool (see Varian Medical Systems, Inc [2015b]) and they can be assumed to be clinically realistic. The plans were optimized until the optimizer algorithm converged according to its default convergence criteria. These two plans (denoted A0 and B0 hereafter) were then used as the reference for the following experiments.

Four variations were created from both base plans. First replaces the original objectives with DVH line objectives (upper for OARs, exact for targets) created from the DVH curves of the reference plans A0 and B0. For PTV(s), also lower and upper point objectives were used, based on doses at 2% and 98% volumes. The objectives were set to 0% and 100% volumes respectively. The stretching was done in order to convey the original wish better to the optimizer. The resulting plans (denoted A1 and B1), are used as a baseline against which the spatial objectives are compared. Using the dose-volume objectives is one possible heuristic for plan conversion. For example the plan conversion method currently used in Eclipse is based on DVHs. Here it also allows to compare the differences in the resulting dose as well as the time the optimization takes.

In the second variation the original objectives were replaced with lower and upper spatial objectives. Both spatial objectives used the dose of the reference plans (A0 or B0), and they had equal priorities, creating in effect an exact objective that aims to reach the reference dose. These plans are denoted by A2 and B2.

The third set of plans (denoted A3 and B3) is a variation of the second, such that the lower spatial objective is only set for a control structure (denoted PTV+10 mm), the volume of which consists of the the target structure(s) extended with a 10 mm margin. In theory having a lower objective for the whole body should not worsen the results if the optimizer is able to reach the wanted dose exactly. However, by definition the PTV includes all the volume containing tumor, and hence there is no clinical reason to have a lower limit on the dose outside of the PTV. Especially if the optimizer’s dose matrix does not match the spatial objective matrix and interpolation is needed, the interpolated lower objective values might cause the optimizer to increase the dose to other body parts compared to the original plan. Plans A3 and B3 are compared to A2 and B2 to analyze these effects. The margin size of 10 mm was chosen based on informal tests where it seemed to produce best results. For these test cases the optimizer dose matrix has a resolution of 1.875 mm and a 10 mm margin is therefore equivalent to approximately five voxels. The optimal size of the margin is studied in more detail in Section 5.3.5.

The fourth set of plans is a combination of variations 1 and 2: exact spatial objectives and dose-volume objectives generated from the reference plan’s DVHs. The priorities of the spatial objectives were set such that their costs are of the same magnitude as the sum of costs of the DVH objectives. These plans are denoted by A4 and B4.

The plans were then optimized for different number of iterations. The used iteration counts were “auto”, 50, 100, 200, and 1000 for case A and “auto”, 20, 35, 50, and 1000 for case B. Here “auto” indicates that the optimization was stopped once it detected convergence. The 1000 iterations versions were done in order to see if the local optimum effect also occurs in the cases with realistic dose distributions. Otherwise only the iteration counts close to the “auto” convergence are considered, as optimizing for 1000 iterations is not feasible in practice. The iteration count is shown as a suffix, e.g., A1-50.

5.3.2 The metrics used to analyze results

The comparisons here are based on the speed (time per iteration and number of iterations required) and on how close the result is to the reference plan’s dose. The accuracy is measured using the root mean square error

$$\text{RMSE}(S) = \sqrt{\frac{1}{|S|} \sum_{i \in S} (x_i - x_i^o)^2}, \quad (14)$$

where x_i^o is the dose value at voxel i in the reference dose, x_i is the dose value at voxel i in the dose for which the RMSE is calculated, S is the set of voxels for which the RMSE is calculated, and $|S|$ is the number of voxels in the set. The dose values are relative to the prescribed dose, so the unit of the RMSE is percentage points, abbreviated as *pp*. For each variation the RMSE is calculated for all relevant structures, including the body structure which contains the whole dose matrix.

It is worth noting that the RMSE is not related to the clinical quality of a plan. For example, a plan having lower dose to an OAR when compared to the reference plan leads to that structure having a large RMSE, even if the lower dose would be clinically favorable. However, the aim of this section is to evaluate if the spatial objectives can be used to produce a given dose distribution, and for that purpose the RMSE is relevant. The clinical aspect is taken into consideration in the plan conversion case studied in Section 6.

5.3.3 Results

The results in Table 1 show that the spatial objectives are capable of directing the optimizer very close to a desired dose distribution at least when it is physically feasible.

On all test cases, the versions utilizing spatial objectives reach a better (in terms of lower RMSE for body) result than the one based on only DVH curves. For the PTV+10mm structures no DVH objectives were generated, as those are not clinical structures but only used for control purposes, e.g., to set the spatial objectives and to calculate the RMSE. In the case with only DVH-objectives (A1) the control structure has the highest RMSE of all structures, while using spatial objectives it

Table 1 – The RMSEs for the body structure (i.e., the whole dose matrix) for different iteration counts. Variation 1 has the highest RMSE of the four cases, while the versions that use spatial objectives have significantly lower RMSE.

	1 [pp]	2 [pp]	3 [pp]	4 [pp]
A-50	1.446	0.215	0.212	0.183
A-100	1.423	0.197	0.178	0.145
A-200	1.423	0.191	0.172	0.105
B-20	4.692	0.426	0.569	0.410
B-35	4.591	0.380	0.481	0.345
B-50	4.563	0.371	0.458	0.342

has RMSE comparable to other structures (c.f. Table 2). This suggests that it is important to specify objectives for all relevant structures when using the DVHs as means of replicating a dose.

Table 2 – The RMSEs for the different structures and variations for case A.

Plan	Body [pp]	PTV [pp]	PTV+10mm [pp]	Bladder [pp]	Rectum [pp]
A1-200	1.423	0.572	2.758	1.96	2.005
A2-200	0.191	0.654	0.569	0.368	0.481
A3-200	0.172	0.575	0.507	0.330	0.479
A4-200	0.105	0.274	0.293	0.214	0.210

For the PTV, the DVH objective results in a lower RMSE than the spatial objectives. This is probably due to the PTV point objectives being tightened so that they ended up being quite close to the original PTV objectives. This in addition to the PTV objectives having the highest weights may lead the optimizer to produce similar dose for the targets. Also the difference in the objective weights explains why the OARs show larger error than the PTV for variation 1.

The difference in the results when setting the lower spatial objectives to the whole body or only for the target is small (variations 2 and 3). Analyzing the target DVHs in the two variations (Figure 10), both versions seem to have similar behavior (smoother corners meaning less homogeneous dose), but for the variation 3 the mean dose is much closer to the reference plan, resulting in slightly smaller RMSE.

The results after 1000 iterations are shown in Table 3. When using only dose-volume objectives, there is very little or no improvement in the body RMSE. When spatial objectives are used, the improvement is significant. An exception is the B3-variation, where the additional iterations do improve the result slightly, but not that much. The result is due to the optimizer not increasing the line search range sufficiently in order to find a better solution. These results again suggests that there is room for improvement in the optimizer when the spatial objectives are used.

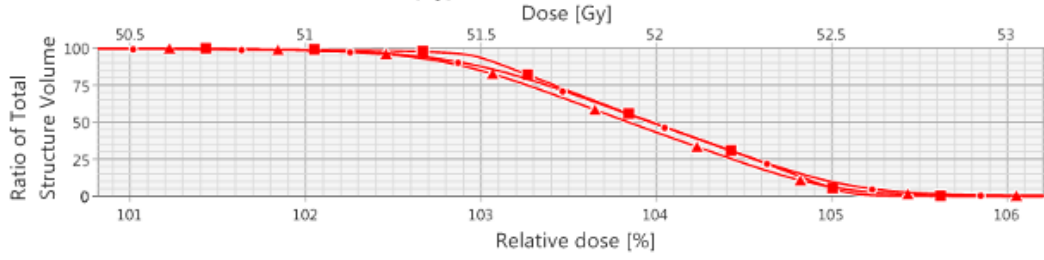


Figure 10 – The DVH curves for the PTV in the reference plan (squares), variation 2 (triangles) and variation 3 (circles). Both variations lead to similar curve, but the variation 2 has slightly worse mean dose.

Table 3 – The RMSEs for the body structure (i.e., the whole dose matrix) at convergence and after 1000 iterations. For case A also the results after 200 are shown. For the “auto” case, the number of iterations is shown in brackets.

	1 [pp]		2 [pp]		3 [pp]		4 [pp]	
A-auto	1.525	(36)	0.200	(82)	0.180	(93)	0.135	(122)
A-200	1.423		0.191		0.172		0.105	
A-1000	1.423		0.006		0.000		0.021	
B-auto	4.527	(90)	0.373	(44)	0.463	(45)	0.320	(55)
B-1000	4.488		0.028		0.337		0.046	

5.3.4 Effect of spatial objectives in iteration time

The speed of optimization depends on two factors: the time spent per iteration and the number of iterations required. Not much can be said from the required number of iterations due to the small amount of data used here, but the time spent per iteration can be analyzed.

Table 4 – Times per iteration for each variation of objectives for cases A and B. The times were averaged over five runs, and the standard deviations are also shown. The optimizer is deterministic, so the variation comes from the computer on which the tests were run. The time reported by the optimizer includes some preprocessing overhead, leading to systematic difference in time spent per iteration depending on the number of iterations. Here only the comparison between the variations is of interest.

Variation	0 [ms]	1 [ms]	2 [ms]	3 [ms]	4 [ms]
A100:	513 \pm 10	525 \pm 6	677 \pm 17	637 \pm 14	672 \pm 18
B50:	464 \pm 36	477 \pm 14	535 \pm 11	534 \pm 8	599 \pm 20

For both patients, the original objectives and the copied DVH objectives result in fastest iterations (Table 4). Using the spatial objectives is slower. Restricting the lower spatial objective to the target makes the iterations slightly speedier since the voxels where the objective is not set can be skipped faster.

Using many DVH objectives makes iterations slower. Each structure that has at least one volumetric constraint (DVH point or line objective) needs to have its DVH calculated. In addition, the calculation time of the costs for a structure is relative to the volume of the structure and the number of dose-volume objectives it has. The spatial objectives' cost calculation time is relative to the volume of the objective matrix. However, most of the time during one iteration is spent in other parts of the optimizer, so the differences caused by the objectives are fairly small.

The time the optimizer uses per iteration is a sum of many different operations. A break-down of an average iteration is shown in Table 5. Approximately one third of the time is spent on calculating the gradient matrix, projecting it to fluence gradients, and calculating the dose change. The rest is used in the line search phase and to send and receive data between the optimizer and Eclipse.

Calculating the gradient for the dose matrix is quite fast (order of 10 ms), but it reflects the computational costs of the objectives quite well. Variation 1, that uses only dose-volume objectives, is faster than the variations 2 and 3 that use spatial objectives. The variation 4 that combines DVH objectives and spatial objectives is the slowest. A notable difference is the field gradient calculation phase, which seems to be much slower when spatial objectives are used. The reason is probably that when only dose-volume objectives are used, large amount of the voxels in the gradient matrix are zero and can be skipped. The spatial objectives result in most voxels in the gradient matrix having a value different from zero. Also the time taken by line search, input, output, and other functions is significantly higher for the versions that use spatial objectives. That is caused at least partly due to inefficient handling of input data (i.e., some properties of the spatial objective matrices are read on each iteration).

Table 5 – Time consumption of different steps in optimization for an average iteration. The times are calculated for a representative run from the test summarized in Table 4. The times are in milliseconds. Also the percentages of total time spent are shown.

	A0 [ms]	A1 [ms]	A2 [ms]	A3 [ms]	A4 [ms]
Dose gradient	10 (3%)	7 (2%)	15 (3%)	13 (3%)	18 (3%)
Field gradients	64 (17%)	65 (17%)	96 (19%)	82 (17%)	96 (19%)
Doses	55 (15%)	56 (15%)	55 (11%)	54 (11%)	55 (11%)
I/O and others	239 (65%)	250 (66%)	352 (68%)	335 (69%)	340 (67%)
TOTAL	369	378	517	485	509

5.3.5 Choosing an optimal margin

Based on the two test cases in Table 6, it is not clear whether it is beneficial to have the lower objective for full body. In case A, setting the lower spatial objective only to the target produces better results for each structure, including body and PTV. It

also improves the speed. In case B, the all-body version produces lower RMSE for body but higher for both targets. Most of the OAR structures have lower RMSEs on the full-body version, but it can be argued that the PTV dose is the most important. The differences are nevertheless very small.

Table 6 – The RMSEs for different structures and the two ways of using spatial objectives for cases A and B. In case A, OARs 1 and 2 are bladder and rectum, respectively. In case B, the OARs are brainstem and spinal cord, and the PTV marks the union of both PTVs.

Plan	Body [pp]	PTV [pp]	OAR 1 [pp]	OAR 2 [pp]
A2-200	0.191	0.654	0.368	0.481
A3-200	0.172	0.575	0.330	0.479
B2-50	0.371	1.322	0.276	0.371
B3-50	0.458	1.292	0.235	0.533

In order to test there is an optimal margin for the lower spatial objective, the cases A3 and B3 were optimized with spatial objectives such that the lower objective was set for different volumes. The volumes used were whole body, PTV, and PTV with different margins (-10 mm, 5 mm, 10 mm, 20 mm, 30 mm, and 50 mm). The RMSEs were calculated for the PTVs and OARs. The negative margin produced very bad results in terms of the RMSE, as one could expect. The optimizer tries to increase the dose level only in the voxels for which a lower objective has been specified. Thus, not specifying a lower dose objective for the whole PTV leads to improved OAR sparing but clinically unacceptable dose coverage for the PTV.

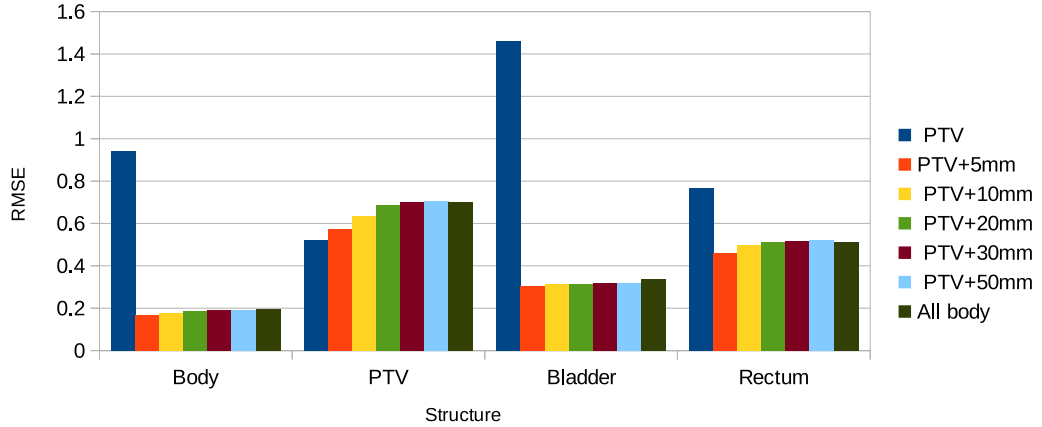


Figure 11 – The RMSEs for case A after 100 iterations of optimization. The lower objective is set for seven different volumes: PTV, PTV with different margins, and full body. Different volumes are represented with different colors.

The results for non-negative margins are shown in Figures 11 (case A) and 12 (case B). All variants with at least some margin produced similar results, while the PTV-only

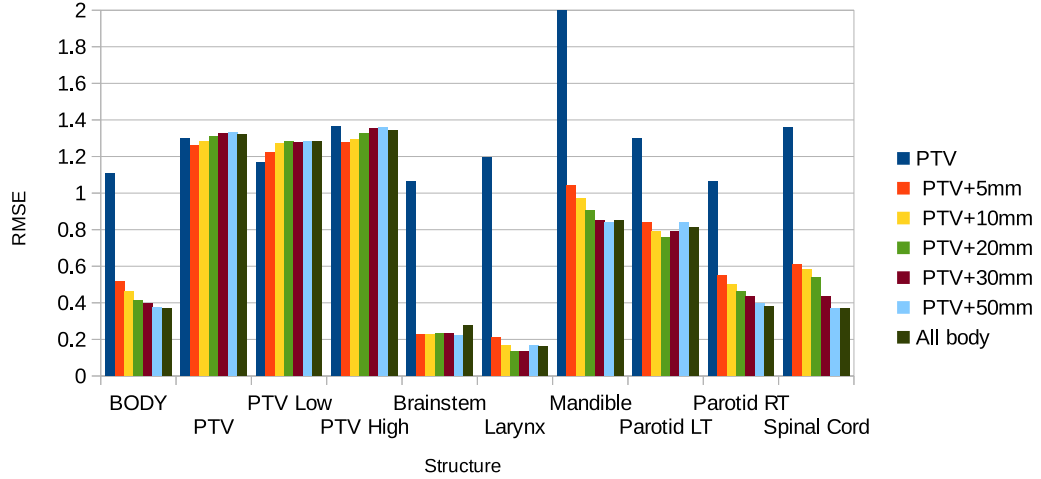


Figure 12 – The RMSEs for case B after 50 iterations of optimization. The lower objective is set for seven different volumes: PTV, PTV with different margins, and full body. Here PTV means the union of PTV Low and PTV High. Different margins are represented with different colors.

version produced relatively bad results on almost all structures for both patients. As an exception, the PTV-only version produced the best result for PTV. The result may be due to a smaller volume of exact spatial objective being able to produce a better result inside the volume where it is specified as it leaves the optimizer more freedom outside the exact objective.

A close-up of the PTV DVHs for case A reference plan, the PTV-only variant, and the 10 mm-margin variant are shown in Figure 13. While the PTV-only variant results in a DVH that is closer to the reference plan’s DVH for the dose fall-off, the 10 mm-margin case has smaller low-dose volume. Having a less steep slope in the PTV-only case allows the optimizer to reduce the dose to other parts of the body, thus leading to a high RMSE for other structures.

Based on the obtained results it seems that setting the lower spatial objective for the PTV only is not enough. Instead, a margin of at least 5 mm should be used around the target in order to prevent underdosage in the border of the target. The effect of the size of the margin seems to be insignificant. The lower spatial objective can also be set to the whole body, at least in these cases where the targeted dose distribution is known to be possible. For case A, a 5 mm margin seems to be optimal, while for patient B setting the lower spatial objective to the whole body results in the smallest body RMSE. It is likely that the optimal margin has to be determined on case-by-case basis. Nevertheless, the differences are small.

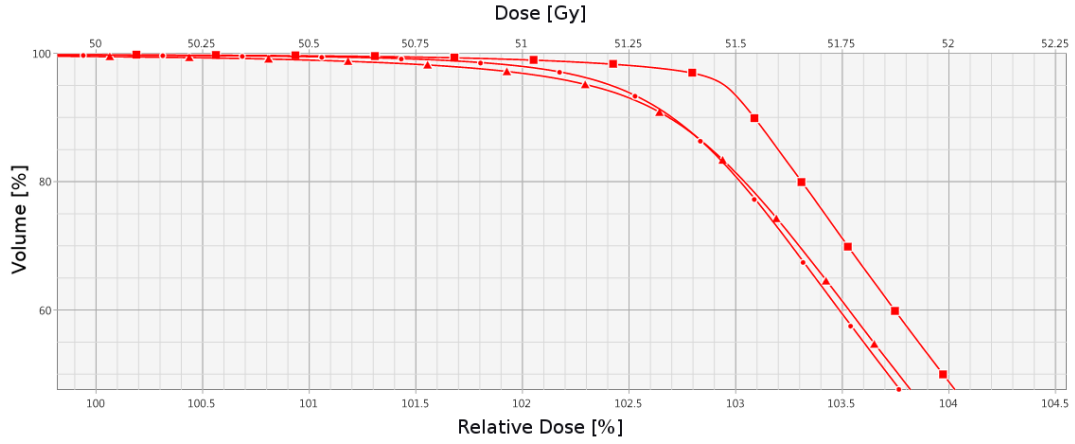


Figure 13 – The PTV dose volume histogram for the case A reference plan (squares), for the PTV-only variant (triangles), and for the PTV plus 10 mm-margin variant (circles). The PTV-only variant provides a DVH closer to the reference plan but the 10 mm-margin variant has better dose homogeneity. Overall the differences are very small.

5.3.6 Choosing an optimal priority

In order to test the effect of the (relative) priorities of the objectives plans for cases A3 and B3 were optimized with different priorities for the lower spatial objective, while keeping the upper spatial objective priority constant (100). The priorities used for the lower spatial objective were 50, 75, 90, 100, 110, 125, 150 and 200.

The results for the case A are shown in Figure 14. Although the differences are not very large, the results indicate that the priority of the lower spatial objective should be at least as large as the upper spatial objective priority. The best results are obtained with priorities 125 (PTV) and 200 (Body).

For the case B, the results are similar (see Figure 15). Again the best priority depends on the structures, but for the PTVs and most OARs it seems that having a slightly stronger lower objective (100-125) as compared to the upper objective is beneficial. Again, the differences are small and the optimal priorities need to be determined on a case-by-case basis.

The weights of the dose-volume objectives on a structure are normalized by dividing them with the volume of the structure. As the cost of a dose-volume objective (Eq. (5) in Section 4.2.1) is summed over all voxels of the structure, the normalization is needed to render the objective's apparent weight independent of the structure's volume. The weights of the spatial objectives are not normalized. Normalizing the weights might help in finding priorities that work for most cases. However, it also makes the spatial objectives' weights depend on the volume that receives little or no dose.

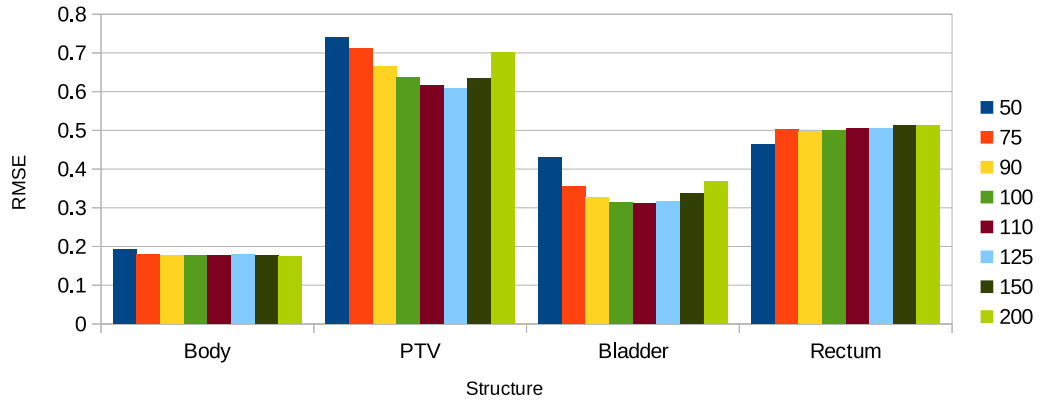


Figure 14 – The RMSEs for case A after 100 iterations of optimization for varying lower spatial objective priorities. The lower objective is set for PTV+10mm margin. The different priorities are represented with different colors (cf. legend on the right).

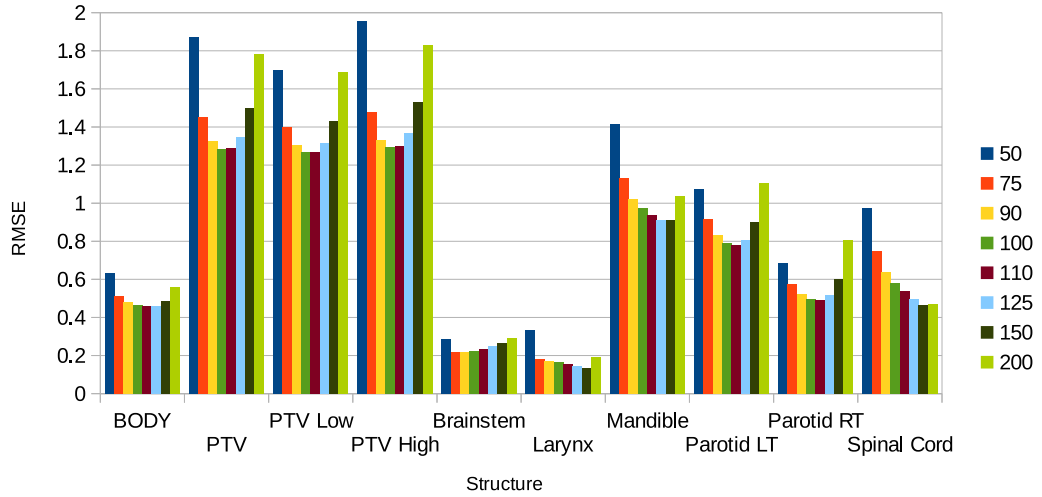


Figure 15 – The RMSEs for case B after 50 iterations of optimization for varying lower spatial objective priorities. The lower objective is set for PTV + 10mm margin. Here PTV denotes the union of PTV Low and PTV High. The different priorities are represented with different colors (cf. legend on the right).

6 Plan conversion using spatial objectives

In the plan conversion use case the aim is to recreate an existing dose distribution on another treatment machine. As the machines may be dosimetrically different, the best result is achieved by re-optimizing the plan for the new machine.

In this Section the spatial objectives will be evaluated as a method for performing plan conversion. The resulting plans will be compared to the original plans and those produced by the *DVH based plan converter* (Varian Medical Systems, Inc [2015b]) currently used in Eclipse (later referred to as *Current*). The quality of the converted plans is analyzed using the criteria discussed in Section 3.

6.1 Test set of treatment plans

A set of 18 clinical plans provided by Varian was used for testing the plan conversion. Each plan corresponds to a different patient. Varian’s clinical personnel has confirmed that the plans are clinically acceptable. Ten of the patients are prostate cases and eight are head and neck (HN) cases.

In all prostate cases (P1 - P10) the clinically relevant structures are a single PTV, bladder, rectum, and femoral heads (left and right). The prescriptions range from 66.6 Gy to 79.2 Gy.

The head and neck cases (HN1 - HN8) are more complex, consisting of one to three targets, and many OARs. The targets have prescriptions ranging from 52 Gy to 70 Gy.

The original plans have been optimized for Varian’s TrueBeam or C-series machines using a 120 leaf MLC (Millennium 120). The prostate plans use energy of 18 MV and the head and neck cases use 6 MV.

6.2 Test setup

A Varian C-series clinac is used as the target machine to which the plans will be converted. It is capable of delivering an IMRT treatment with the same beam configuration as defined in the original plan. The beam configuration has a large impact on the spatial distribution of the dose, so a differing beam configuration would need special handling with regards to the spatial objective. Moreover, in most cases the target machine can provide the exactly same beam configuration as the original machine. Thus only cases where beam configuration remains constant are considered here.

To guarantee that there are differences between the original machine and the target machine, an 80 leaf MLC (Millennium 80) is used for the target machine. Compared to Millennium 120 the Millennium 80 has lower resolution which restricts the modulation

it is capable of providing. Thus, the target machine cannot produce exactly the original dose. For all converted plans the energy is set to 6 MV, producing dosimetric variation in the prostate case. Beam, couch, and collimator angles were kept the same but the converters were allowed to change the field sizes.

The DVH based plan converter tool was used to convert each plan for the target machine in order to create a set of benchmark plans. Only structures deemed clinically relevant were selected to be included in the conversion. The OARs were set as avoidance structures in Eclipse, and the target(s) were marked as PTV structures. The number of fractions for the converted plan was set to match with the original plan so that the dose levels are comparable. A stopwatch was used to time the actual conversion, i.e., excluding the structure selection and the final dose calculation.

A new procedure that utilizes spatial objectives for plan conversion was created based on the results in Section 5. The procedure converts plans using an ESAPI script and a few manual steps in Eclipse. The manual part was required for the parameters that are currently not exposed through the ESAPI. This conversion method is later referred to as *Spatial*s. The plan conversion works as follows

1. Create a copy of a source plan with the target treatment machine and beams correctly set up.
2. Set the upper spatial objective (corresponding to dose from the source plan) for whole body.
3. Set the lower spatial objective (corresponding to dose from the source plan) for the volume covered by targets extended by a 10 mm margin.
4. Optimize with automatic stop on convergence.
5. Perform leaf motion calculation and dose calculation.
6. Continue optimization with the achieved dose as intermediate dose, using again automatic stop on convergence.
7. Perform final leaf motion calculation and dose calculation.

For the lower spatial objective, a priority of 165 was used, while the upper objective's priority was set to 100. The relative priorities of the spatial objectives used here differ from the ones suggested in Section 5.3.6. Having only slightly more priority to the lower objective works well when the dose distribution is reachable. Here, however, the exact dose cannot be reached due to the target machine having an inferior MLC when compared to the original machine, and the whole-body upper spatial objective will restrict the dose quite heavily. To prevent underdosage, the lower spatial objective needs higher priority.

The optimizer algorithm uses a faster and less accurate dose calculation than the final dose calculation, which results in a difference between the dose received produced the optimizer and the one from the final dose calculation. As explained in Section 4.1, the difference can be reduced by using an intermediate dose calculation. Here especially

the leaf motion calculation seemed to affect the resulting dose and doing a full leaf motion calculation and dose calculation as the intermediate dose seemed to improve the results. In some cases the intermediate dose also helps to improve the dose homogeneity for the target and to reduce the OAR doses.

First tests of the conversion showed very high dose hot spots in some HN cases. The reason was that the volume for which the lower objective was set went over the body volume. To reduce the hot spots, the margin in the volume for which the lower objective was set was retracted by 2 mm from the body outline. The whole target volume was still included in the lower objective volume.

After the converted plans were generated, they were normalized such that the PTV D_{mean} is the same for the original and the two converted plans. For the plans that have multiple PTVs the one with highest dose prescription was used for the normalization. The normalization was done in order to make the plans comparable. Different normalization point would yield slightly different results.

Another ESAPI script was used to collect predefined metrics from each plan. For prostate and HN cases different dose-volume metrics were specified based on Emami [2013]. The metrics are meant to indicate the clinically relevant points for comparison. The main criteria used is D_{max} for serial organs and D_{mean} for parallel organs. In addition to D_{max} also $D_{2\%}$ is reported according to the ICRU83 recommendation.

The original plans have aimed for a homogeneous dose to targets. To see if the homogeneity is preserved in the conversion, the homogeneity index [see Eq. (2), Section 3.4] was calculated for each target.

Additional criteria include D_{max} and D_{mean} for the body structure to confirm that the overall dose level stays approximately the same and to make sure that no points with very high dose are produced outside the targets. Also the HOT measure (Section 3.4) is reported. The RMSE metric is calculated for all voxels where the dose is significant ($> 10\%$ of prescribed dose) in the converted plan. The low-dose regions are mostly scattered dose and can be assumed to be similar in the original and in the converted plans.

6.3 Plan conversion for prostate cases

The metrics that were used for the OARs in the prostate cases are listed in Table 7. For the PTVs, the near-min, near-max, and max doses are reported, as well as the homogeneity index. D_{mean} and D_{max} were also calculated for the body structure, as well as the hot spots outside target volume and the root mean square error. In all of the converted plans the overall maximum dose was inside the PTV, so the body D_{max} is not reported.

A summary of the results is shown in Table 8. The average values of the metrics in the original plans are shown for reference. For the converted plans, the mean differences are shown, as well as the maximum differences. For PTV $D_{98\%}$ a high

Table 7 – Dose-volume metrics for the prostate cases, based on Emami [2013]. The goals are indicative only.

Organ at Risk	Metric	Goal
Femoral Heads	$D_{max}[Gy]$	52
Bladder	$V_{65Gy}[\%]$	50
Bladder	$V_{70Gy}[\%]$	35
Rectum	$V_{50Gy}[\%]$	50
Rectum	$V_{65Gy}[\%]$	35

value is preferable, but for all other metrics a lower value is better. The converted plans have been normalized to have the same D_{mean} for the PTV as in the original plan. The average amounts of normalization were $1.12\% \pm 1.01\%$ and $0.60\% \pm 0.26\%$ (i.e., the dose levels were scaled up by about 1%) for the plans converted with the *Spatials* and *Current* methods, respectively.

Table 8 – The results of the plan conversions for the prostate cases. First the metrics’ average values and standard deviations for the original plans are shown. For the converted plans, the average and maximum differences are shown. For PTV $D_{98\%}$ a high value is preferable, otherwise a lower value is better. Fem. Head, L and Fem. Head, R stand for left and right femoral heads, respectively.

Structure	Metric	Original		Δ Spatials		Δ Current	
		Mean	σ	Mean	max	Mean	max
PTV	$D_{98\%}[Gy]$	73.80	6.25	-1.63	-4.23	-1.63	-2.50
PTV	$D_{2\%}[Gy]$	76.88	6.39	1.10	3.65	0.52	0.97
PTV	$D_{max}[Gy]$	77.64	6.01	2.58	7.34	1.40	2.55
PTV	HI	0.04	0.01	0.04	0.06	0.03	0.05
Fem. Head, L	$D_{max}[Gy]$	46.13	4.55	2.02	9.14	1.97	12.27
Fem. Head, R	$D_{max}[Gy]$	41.92	7.41	-2.47	-0.57	0.76	5.71
Bladder	$V_{65Gy}[\%]$	21.62	18.06	-0.01	0.89	-0.55	0.78
Bladder	$V_{70Gy}[\%]$	5.83	9.32	0.22	1.69	0.31	2.97
Rectum	$V_{50Gy}[\%]$	20.19	10.94	1.10	1.90	2.40	5.55
Rectum	$V_{65Gy}[\%]$	10.92	5.75	-0.23	0.53	0.01	0.78
Body	$D_{mean}[Gy]$	4.05	1.38	0.34	0.46	0.50	0.83
	HOT [cc]	18.25	8.68	-3.39	-1.01	7.86	22.84
	RMSE [pp]			5.82	7.60	8.18	10.32

Compared to the original plans, the *Spatials* conversion produces slightly worse plans. This is expected, as the target machine is not capable of producing the dose distribution of the source plan. On average, the conversion results in significantly worse PTV dose homogeneity that is visible in the homogeneity index and PTV’s $D_{2\%}$, $D_{98\%}$ and D_{max} . The average difference is also significant for left femoral head’s D_{max} , where the converted plans are worse than the originals. Interestingly, the

converted plans have smaller HOT volume than the original plans. This may be related to having also smaller PTV $D_{98\%}$. The definition of HOT is the volume outside target with $> 100\%$ dose. Having a slightly reduced $D_{98\%}$ for the PTV allows the optimizer to reduce the dose in other parts of the body below the prescription dose, thus reducing the HOT volume. For other metrics, the average differences are small or show improvement in the converted plan.

The maximum differences in the metrics between the plans converted with the *Spatials* method indicate the same behavior. The conversion has problems with the PTV dose homogeneity, resulting in very $D_{98\%}$, $D_{2\%}$ and D_{max} for the PTV in some cases.

The metrics calculated from the converted plans were also compared against the goals in Table 7. Mostly the goals were achieved. The $V_{65Gy} \leq 50\%$ goal was not achieved in one case, but in that case the volume was actually even higher in the original plan. The goal of $D_{max} \leq 52$ Gy for the left femoral head was exceeded in three cases. In two of those the converted plan had significantly higher dose than the original plan. The conversion made with the *Current* method resulted in even higher dose for the left femoral head in two of the three cases.

Compared to the *Current* method, the *Spatials* method seems to produce similar or better plans in terms of everything else except the PTV near-maximum and maximum doses. Especially the RMSE metric is clearly smaller for the plans converted using the spatial objectives, meaning that the spatial objectives do guide the optimizer to a solution that is on average closer to the original dose distribution. An example comparison of DVHs is shown in Figure 16.

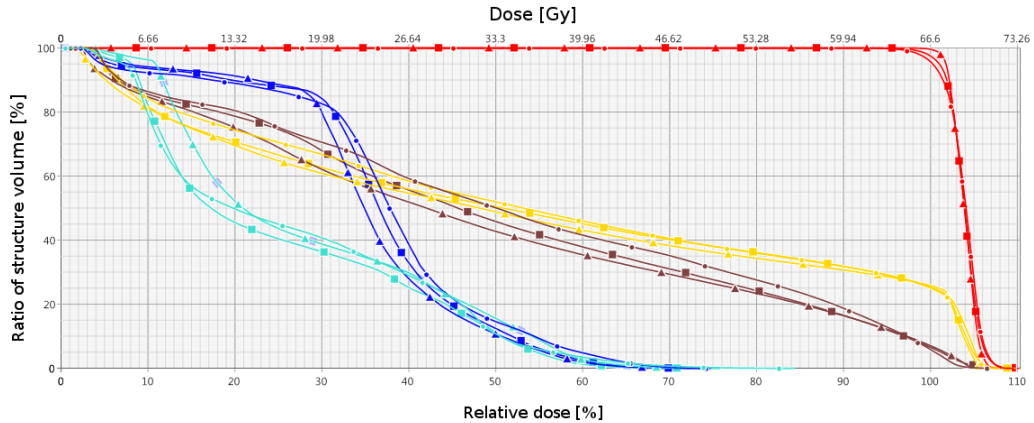


Figure 16 – A comparison of the DVHs for P1. Triangles mark the original plan, squares the *Spatials* conversion and circles the *Current* conversion. PTV is red, blue and cyan are left and right femoral heads, and yellow and brown are for bladder and rectum.

Unlike the *Spatials* method, the *Current* method results in a significant increase in the volume of hot spots outside target. This occurs also in cases where the *Spatials* conversion results in higher $D_{98\%}$ for the PTV. Since the DVH based plan converter only considers the structures selected for the conversion (in this case the PTV, OARs

and body), it allows higher doses in the tissue not covered by the PTV or OARs. The upper spatial objective prevents the dose from increasing in any part of the body, leading to less hot spots in the *Spatial*s method.

Three of the ten cases are shown in Appendix A. Tables A1 and A3 show cases where the *Spatial*s conversion produced overall better plans than the *Current* method. A case where the *Spatial*s method resulted in significantly worse plan (due to low PTV homogeneity) is shown in Table A2.

To summarize, it seems that the conversion method based on the spatial objectives produces overall more similar dose distribution, but has problems with PTV’s near-minimum, near-maximum and maximum doses. Excluding the PTV homogeneity the plans produced with the *Spatial*s conversion are, according to the criteria used here, of at least as high clinical quality as the plans produced with the *Current* conversion method.

In addition to the quality of the conversion, the time it takes is also of interest. The average times for the prostate cases are shown in Table 9. Excluding the final dose calculation (that is common to both methods) and the structure selection step (required by the *Current* method), the *Spatial*s method takes almost twice as long as the *Current* method. Still, the *Spatial*s method takes approximately 2 min which should be acceptable in practice. It should be noted that the *Spatial*s method is implemented through ESAPI. Implementing it natively as a part of Eclipse would likely make it faster. There is also much room for optimization in the conversion process if speed is deemed important.

The *Spatial*s method uses less than half of the time for optimization. The rest goes for setting up the conversion and performing the intermediate dose calculation. Without the intermediate dose calculation and second optimization, the *Spatial*s conversion would be faster than the *Current* method, but the quality of the conversion would suffer.

Table 9 – The running time of the converters, averaged over the ten prostate cases. The first row corresponds to the DVH based converter currently used in Eclipse. The second row is the time the *Spatial*s conversion takes. The last four rows show a break-down of the *Spatial*s conversion. Also standard deviations are shown.

	Mean [s]	Iterations
<i>Current</i>	60 ± 20	
<i>Spatial</i> s	105 ± 14	
Setup	31 ± 6	
1st run	23 ± 4	20 ± 2
Dose calc	31 ± 4	
2nd run	20 ± 4	11 ± 0

6.4 Plan conversion for head and neck cases

The metrics that were used for the HN patients' OARs are listed in Table 10. For the PTVs, the near-min, near-max, and max doses are reported, as well as the homogeneity index. The plans had one to three targets with differing dose levels. Here PTV High denotes for each plan the target with the highest dose, PTV Med the target with the second highest dose (if the plan has two or three PTVs), and PTV Low denotes the target with the lowest dose prescription (if the plan has three PTVs). The converted plans have been normalized to have the same D_{mean} for the PTV High as in the original plan. For PTV Med the D_{mean} is also reported. Only one case had three PTVs, so the third one (PTV Low) is not listed in the averaged results. For the body structure D_{mean} , D_{max} , and hot spots outside target were calculated. Also the root mean square error is reported. The average amounts of normalization were $-1.04\% \pm 0.36\%$ and $-1.14\% \pm 0.62\%$ (i.e., the dose levels were scaled down by about 1%) for the plans converted with the *Spatial*s and *Current* methods, respectively.

Table 10 – Dose-volume metrics for the head and neck cases, based on Emami [2013]. The goals are indicative only.

Organ at Risk	Metric	Goal
Brain	$D_{max}[Gy]$	60
Brainstem	$D_{max}[Gy]$	64
Chiasm	$D_{max}[Gy]$	50
Ear	$D_{mean}[Gy]$	45
Esophagus	$D_{mean}[Gy]$	34
Eye	$D_{max}[Gy]$	50
Larynx	$D_{mean}[Gy]$	44
Mandible	$D_{max}[Gy]$	70
Optic Nerve	$D_{max}[Gy]$	50
Parotid	$D_{mean}[Gy]$	20
Spinal Cord	$D_{max}[Gy]$	50

The results for the head and neck cases are summarized in Table 11. Again the plans converted with the *Spatial*s method have mostly worse metrics than the original ones, as is to be expected. On average the plans converted with the *Spatial*s have slightly worse PTV dose homogeneity on both the high-dose PTV and the medium-dose PTV, resulting in a lower $D_{98\%}$ and a higher $D_{2\%}$, but the difference is smaller than in the prostate cases. The D_{max} of the high dose PTV is still on average significantly higher than in the original plans.

In one case, the PTV High's D_{max} is almost 10 Gy above the original plan, and one case resulted in 5 Gy higher D_{max} . In both cases there is a hot spot inside the PTV High close to the body surface. This suggests that the lower spatial objective causes the optimizer to try to increase the dose to the voxels near the body outline, but due

Table 11 – The results of the plan conversions of the prostate cases. First the metrics’ average values and standard deviations for the original plans are shown. For the converted plans, the average and maximum differences are shown. For PTV $D_{98\%}$ a high value is preferable, otherwise a lower value is better. For chiasm, eyes and optic nerves the doses were insignificant (< 10 Gy) and thus they are not shown here. PTV High means for each patient the PTV with the highest dose prescription, and PTV Med the second highest. Only one patient had three PTVs so the third one (PTV Low) is not included.

Structure	Metric	Original		Δ Spatials		Δ Current	
		Mean	σ	Mean	max	Mean	max
PTV High	$D_{98\%}[Gy]$	62.74	4.89	-0.49	-0.87	-0.97	-1.35
PTV High	$D_{2\%}[Gy]$	68.81	5.41	0.80	1.26	0.83	1.11
PTV High	$D_{max}[Gy]$	71.00	5.58	3.76	9.89	1.18	2.63
PTV High	HI	0.09	0.02	0.02	0.03	0.03	0.04
PTV Med	$D_{98\%}[Gy]$	55.70	7.49	-0.56	-0.92	-0.61	-1.71
PTV Med	$D_{2\%}[Gy]$	61.93	8.24	0.71	1.22	0.99	1.68
PTV Med	$D_{max}[Gy]$	59.08	7.72	0.06	0.23	0.43	0.85
PTV Med	$D_{mean}[Gy]$	64.67	8.69	1.62	2.93	1.90	3.99
PTV Med	HI	0.11	0.03	0.02	0.04	0.03	0.05
Brain	$D_{2\%}[Gy]$	20.17	17.71	-0.34	0.95	1.60	6.10
Brain	$D_{max}[Gy]$	38.21	25.81	0.07	2.95	0.48	4.29
Brainstem	$D_{2\%}[Gy]$	19.29	15.24	-0.11	1.54	-0.47	1.36
Brainstem	$D_{max}[Gy]$	23.01	16.98	0.07	3.45	0.57	5.02
Ear, L	$D_{mean}[Gy]$	10.09	11.76	0.82	1.95	-0.93	0.13
Ear, R	$D_{mean}[Gy]$	5.15	3.72	0.83	1.50	0.24	1.29
Esophagus	$D_{mean}[Gy]$	34.49	16.60	0.62	2.45	0.30	2.01
Larynx	$D_{mean}[Gy]$	35.01	19.55	1.21	2.79	0.68	1.71
Mandible	$D_{2\%}[Gy]$	61.49	10.97	0.47	1.28	0.54	3.53
Mandible	$D_{max}[Gy]$	66.31	9.42	1.27	6.03	0.07	4.07
Parotid, L	$D_{mean}[Gy]$	22.12	20.72	0.97	2.24	0.33	0.84
Parotid, R	$D_{mean}[Gy]$	15.00	17.72	1.25	2.66	1.25	2.62
Spinal Cord	$D_{2\%}[Gy]$	35.27	2.35	-0.22	2.35	1.70	4.37
Spinal Cord	$D_{max}[Gy]$	37.22	2.65	0.07	1.54	2.92	8.54
Body	$D_{max}[Gy]$	71.04	5.61	3.72	9.89	1.57	5.77
Body	$D_{mean}[Gy]$	10.46	2.53	-0.63	0.21	-0.28	1.20
Body	HOT [cc]	34.09	34.29	-1.95	1.61	22.47	49.15
	RMSE [pp]	0.00	0.00	3.85	6.01	8.00	10.14

to dose build-up effect it ends up creating a hot spot inside the PTV. To prevent this kind of problems the margin used for the targets when setting the lower spatial objective volume was retracted 2mm from the border of the body structure. The problem may be caused by the 2mm being not enough, or it may happen because all of the target volume was still included in the lower objective.

In these cases the *Spatials* conversion results in better target homogeneity than the *Current* conversion method for both targets. Also the D_{mean} for PTV Med is closer to that of the original plans when using the *Spatials* method. Out of 14 OAR metrics, seven have smaller average difference when using the *Spatials* method and the other seven when using the *Current* method. Notably the *Spatials* plans show smaller increases for brain and spinal cord. The plans converted with the *Spatials* method have similar volume of hot spots outside target when compared to the original plans, while the *Current* conversion increases the volume significantly. The RMSE metric is smaller for the *Spatials* method, meaning that the dose distributions of the plans converted using the spatial objectives are closer to those of the original plans.

The values calculated for the metrics were also compared against the goals in Table 10. The plans created with the *Spatials* conversion stayed inside the goals in all the cases where the original plan was inside the goals. However, in three cases the *Spatials* conversion further deteriorated two or three metrics that were already outside of the clinical goal. These metrics were the D_{max} for mandible and D_{mean} for esophagus and parotids.

The calculated metrics for three of the eight cases are shown in Appendix B. A case where the *Spatials* method resulted in significantly worse plan (due to a hot spot inside the high-dose PTV) is shown in Table B1. Table B2 shows a case where the *Spatials* conversion produced overall better plan than the *Current* method. Table B3 is an example of a case where the conversion worsened a metric that is already outside of the goal.

To summarize, in most cases the *Spatials* conversion seemed to result in a plan that is overall closer to the original plan. In these cases it also results in better target doses, excluding the two cases with hot spots inside the PTV High. The values for the OAR metrics are similar to those of the plans produced with the DVH based plan converter.

Table 12 – The running time of the converters, averaged over the eight HN cases. The first row corresponds to the DVH based converter currently used in Eclipse. The second row is the time the *Spatials* conversion takes. The last four rows show a break-down of the *Spatials* conversion.

	Mean (s)	STD	Iterations
<i>Current</i>	48	13	
<i>Spatials</i>	139	30	
Setup	24	4	
1st run	44	14	33 ± 7
Dose calc	40	9	
2nd run	31	8	15 ± 2

Table 12 shows the conversion times for the head and neck cases. Interestingly the *Current* conversion is faster than in the prostate cases, but the *Spatials* conversion

is significantly slower. Although the difference is now more notable, the *Spatials* conversion still takes only about 2.5 min, which is acceptable in practice.

The increase in running time for the *Spatials* method compared to the prostate cases is mostly due to the optimizer performing more iterations before convergence. For the HN cases there is also more variation in the number of iterations for both the first and the second optimization runs.

6.5 Additional remarks for further research

Other informal tests were done in addition to the results presented here. The usage of dose-volume objectives in addition to the spatial objectives seemed to result in slightly better values for the metrics, and is something that could be studied further. The objectives were set for the targets (an exact line objective and high priority upper and lower point objectives) and organs at risk (upper line objectives). Especially the upper and lower objectives for the targets improved dose homogeneity significantly. While the spatial objectives are able to direct the optimizer towards the original dose, the dose-volume objectives help the optimizer when the spatial objectives cannot be satisfied.

Using the same dose-volume objectives was also tested without spatial objectives. With no objective affecting the rest of the body, the optimizer achieved greatly improved OAR sparing but created hot spots on the normal tissue. The resulting plans differed greatly from the original plans and were clinically unacceptable. This confirms that the role of the spatial objectives is significant even if both spatial and dose-volume objectives are used.

The results reported here were done with a 10 mm margin around the targets for the lower spatial objective. Also 5 mm and 20 mm margins were tested. The differences in the resulting values for the metrics were small. The optimal margin depends on the case and the metric used to evaluate the results.

One of converted plans was optimized for 1000 iterations to see if the local optima effect discussed in Section 5.2 is visible in the plan conversion. The additional iterations did not improve the results, suggesting that the local optima is a problem only in the case where a zero-cost global optimum exists.

7 Conclusions

The aim of this Thesis was to explore the usage of spatial objectives in radiation therapy treatment planning, focusing on the plan conversion use case. The empirical work was divided into two parts: in the first part the behavior of the spatial objectives was analyzed (Section 5) and then the usage of the spatial objectives in the context of plan conversion was tested (Section 6).

The analysis in Section 5 shows that the spatial objectives do work as a means of directing the optimizer towards a wanted dose distribution. Compared to using dose-volume objectives, the spatial objectives resulted in significantly lower RMSE between the targeted and the achieved doses. Combining the spatial objectives with the dose-volume objectives yielded even lower error. Further testing indicated that in order to achieve the best results, the lower objective should be set at least for the target structures plus an additional margin, otherwise the near-border regions of the target will easily get underdosed. When the objective dose was achievable, setting the lower spatial objective for the whole body was found to be plausible. Giving the lower objective slightly higher priority than the upper objective was found to improve the results.

The results of the plan conversion tests indicate that the spatial objectives are useful in the process. In many cases the conversion utilizing spatial objectives produced better results than the converter used currently in Eclipse, and only in a few cases did the conversion produce plans that are significantly worse than the original plan with regards to one or more of the specified criteria. When compared against the DVH based plan conversion, the spatial objectives resulted in all cases in a dose distribution that is overall significantly closer to the dose distribution of the original plan. The conversion was somewhat slower than the current converter, but there is room for improvement in the speed.

In few cases, the conversion using the spatial objectives resulted in clinically unacceptable results. Using for example dose-volume objectives for the targets in addition to the spatial objectives would probably help to improve the target homogeneity. The problem of hot spots close to the body structure's border can probably be solved by adjusting the volume for which the lower spatial objective is set.

The tests performed here raise many new questions. In Section 5 it was observed that when optimizing for an exact dose distribution, the gradient search got stalled and achieved only small improvements after reaching a good solution, until an internal mechanism allowed it to perform longer jumps in further improve the solution. The reason may be that the optimization landscape has a local optimum or a narrow valley through which the gradient search must pass. Further studies of the optimizer could improve the results. Also, using different optimization algorithms instead or in addition to the gradient search might allow the optimizer to escape the possible local minima and reach the global minimum.

Regarding the plan conversion use case, this Thesis considered only IMRT to IMRT conversions, where the beam configuration was kept unchanged. There may be cases where a VMAT plan needs to be converted to IMRT. In some special cases also an IMRT to VMAT conversion might be useful. In these cases the beam geometry is very different between the original plan and the converted plan. In these cases having an upper spatial objective with constant weight for the whole body probably does not work. For example, when converting from VMAT to IMRT, the normal tissue volume radiated by the IMRT beams will inevitably have higher dose when compared to the VMAT plan, causing very high costs for the upper objective and leading to underdosage. A potential solution would be to accompany the objective matrices with matching weight matrices so that the user could specify which structures or other volumes are most important.

The results of the conversion depend on the priorities of the objectives. The priorities used here were selected based on the tests in Section 5 and then adjusted to fit the plan conversion case. In order to prevent underdosage the lower spatial objective had to be given a higher priority relative to the upper spatial objective than warranted by the test cases. The selected priorities seemed to work well enough for the cases tested in Section 6, but they may not be universally applicable. An adaptive optimization changing the priorities of the objectives based on the progress of the optimizer could make the priorities more robust.

This Thesis focused on using the spatial objectives to recreate a realistic dose distribution. Regarding the dose painting method, i.e., prescribing different levels of radiation into different voxels of the target based on e.g., MRI images, further studies are needed.

The results of this thesis can be used to improve the radiation therapy treatment planning and the treatments received by the patients.

References

- L. Arbea, L.I. Ramos, R. Martínez-Monge, M. Moreno, and J. Aristu. Intensity-modulated radiation therapy (imrt) vs. 3d conformal radiotherapy (3dcrt) in locally advanced rectal cancer (larc): dosimetric comparison and clinical implications. *Radiation Oncology*, 5(1):17, 2010.
- S.M. Bentzen. Theragnostic imaging for radiation oncology: dose-painting by numbers. *The lancet oncology*, 6(2):112–117, 2005.
- S.M. Bentzen, L.S. Constone, J.O. Deasy, A. Eisbruch, A. Jackson, L.B. Marks, R.K. Ten Haken, and E.D. Yorke. Quantitative analyses of normal tissue effects in the clinic (quantec): an introduction to the scientific issues. *International Journal of Radiation Oncology* Biology* Physics*, 76(3):S3–S9, 2010.
- T. Bortfeld, J. Bürkelbach, and W. Schlegel. Three-dimensional solution of the inverse problem in conformation radiotherapy. In A. Breit, A. Heuck, P. Lukas, P. Kneschaurek, and M. Mayr, editors, *Tumor Response Monitoring and Treatment Planning*, pages 503–508. Springer Berlin Heidelberg, 1992. ISBN 978-3-540-54783-9. doi: 10.1007/978-3-642-48681-4_82. URL http://dx.doi.org/10.1007/978-3-642-48681-4_82.
- J. Deasy. Multiple local minima in radiotherapy optimization problems with dose–volume constraints. *Medical physics*, 24(7):1157–1161, 1997.
- B. Emami. Tolerance of normal tissue to therapeutic radiation. *Reports of Radiotherapy and Oncology*, 1(1):123–7, 2013. ISSN 2345-3184. URL <http://radioncology.com/en/articles/2782.html>.
- B. Emami, J. Lyman, A. Brown, L. Cola, M. Goitein, J.E. Munzenrider, B. Shank, L.J. Solin, and M. Wesson. Tolerance of normal tissue to therapeutic irradiation. *International Journal of Radiation Oncology* Biology* Physics*, 21(1):109–122, 1991.
- V. Grégoire and T.R. Mackie. State of the art on dose prescription, reporting and recording in intensity-modulated radiation therapy (ICRU report No. 83). *Cancer/Radiothérapie*, 15(6):555–559, 2011.
- E.J. Hall and A.J. Giaccia. *Radiobiology for the Radiologist*. Lippincott Williams & Wilkins, 2006. ISBN 9780781741514.
- E.C. Halperin, L.W. Brady, D.E. Wazer, and C.A. Perez. *Perez and Brady’s Principles and Practice of Radiation Oncology*. Lippincott Williams & Wilkins, 2004.
- B.F. Hasson, D. Michalski, and M.S. Huq. Tumor control probability (tcp). In *Encyclopedia of Radiation Oncology*, pages 921–922. Springer, 2013.

- R.T. Hoppe, R.S. Cox, Z. Fuks, N.M. Price, M.A. Bagshaw, and E.M. Farber. Electron-beam therapy for mycosis fungoides: the stanford university experience. *Cancer treatment reports*, 63(4):691–700, 1979.
- N. Howlader, A.M. Noone, M. Krapcho, J. Garshell, D. Miller, S.F. Altekruse, C.L. Kosary, M. Yu, J. Ruhl, Z. Tatalovich, A. Mariotto, D.R. Lewis, H.S. Chen, E.J. Feuer, and K.A. (eds) Cronin. SEER cancer statistics review, 1975–2012, 2015. URL http://seer.cancer.gov/csr/1975_2012/. Based on November 2014 SEER data submission, posted to the SEER web site, April 2015.
- ICRU International Commission on Radiation Units and Measurements. Prescribing, recording and reporting photon beam therapy (ICRU report No. 50), 1993.
- M. Langer, E.K. Lee, J.O. Deasy, R.L. Rardin, and J.A. Deye. Operations research applied to radiotherapy, an nci-nsf-sponsored workshop february 7–9, 2002. *International Journal of Radiation Oncology* Biology* Physics*, 57(3):762–768, 2003.
- E.K. Lee, T. Fox, and I. Crocker. Integer programming applied to intensity-modulated radiation therapy treatment planning. *Annals of Operations Research*, 119(1-4):165–181, 2003.
- L.B. Marks, E.D. Yorke, A. Jackson, R.K. Ten Haken, L.S. Constine, A. Eisbruch, S.M. Bentzen, J. Nam, and J.O. Deasy. Use of normal tissue complication probability models in the clinic. *International Journal of Radiation Oncology* Biology* Physics*, 76(3):S10–S19, 2010.
- D. Michalski, M.S. Huq, and B.F. Hasson. Normal tissue complication probability (ntcp). In LutherW. Brady and TheodoreE. Yaeger, editors, *Encyclopedia of Radiation Oncology*, pages 560–560. Springer, 2013. ISBN 978-3-540-85513-2. doi: 10.1007/978-3-540-85516-3_666. URL http://dx.doi.org/10.1007/978-3-540-85516-3_666.
- W.T. Moss and J.D. Cox. *Moss’ Radiation Oncology: Rationale, Technique, Results*. Mosby, 1994. ISBN 9780801669408.
- M. Naghavi, H. Wang, R. Lozano, A. Davis, X. Liang, M. Zhou, S.E. Vollset, A.A. Ozgoren, S. Abdalla, F. Abd-Allah, et al. Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990–2013: a systematic analysis for the global burden of disease study 2013. *Lancet*, 385(9963):117–171, 2015.
- A. Niemierko. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Medical physics*, 24(1):103–110, 1997.
- A. Niemierko. A generalized concept of equivalent uniform dose (eud). *Med Phys*, 26(6):1100, 1999.

- R. Orecchia, A. Zurlo, A. Loasses, M. Krenqli, G. Tosi, S. Zurrida, P. Zucali, and U. Veronesi. Particle beam therapy (hadrontherapy): basis for interest and clinical experience. *European Journal of Cancer*, 34(4):459–468, 1998.
- K. Otto. Volumetric modulated arc therapy: Imrt in a single gantry arc. *Medical physics*, 35(1):310–317, 2008.
- H. Paganetti and T. Bortfeld. Proton therapy. In *New Technologies in Radiation Oncology*, pages 345–363. Springer, 2006.
- E.M. Quan, X. Li, Y. Li, X. Wang, R.J. Kudchadker, J.L. Johnson, D.A. Kuban, A.K. Lee, and X. Zhang. A comprehensive comparison of imrt and vmat plan quality for prostate cancer treatment. *International Journal of Radiation Oncology* Biology* Physics*, 83(4):1169–1178, 2012.
- B.W. Stewart and C.P. Wild. World cancer report 2014. lyon, france: International agency for research on cancer. *World Health Organization*, 2014.
- M. Teoh, C.H. Clark, K. Wood, S. Whitaker, and A. Nisbet. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *The British journal of radiology*, 2014.
- L. Tillikainen, H. Helminen, T. Torsti, S. Siljamäki, J. Alakuijala, J. Pyyry, and W. Ulmer. A 3d pencil-beam-based superposition algorithm for photon dose calculation in heterogeneous media. *Physics in medicine and biology*, 53(14):3821, 2008.
- W Ulmer and D Harder. A triple gaussian pencil beam model for photon beam treatment planning. *Zeitschrift für medizinische Physik*, 5(1):25–30, 1995.
- Varian Medical Systems, Inc. Eclipse photon and electron algorithms reference guide, 2015a.
- Varian Medical Systems, Inc. Eclipse photon and electron reference guide, 2015b.
- S Webb. Optimizing the planning of intensity-modulated radiotherapy. *Physics in Medicine and Biology*, 39(12):2229, 1994.
- H.R. Withers, J.M.G. Taylor, and B. Maciejewski. Treatment volume and tissue tolerance. *International Journal of Radiation Oncology* Biology* Physics*, 14(4):751–759, 1988.
- Q. Wu and R. Mohan. Multiple local minima in imrt optimization based on dose-volume criteria. *Medical physics*, 29(7):1514–1527, 2002.

A Plan Conversion tests: Prostate

Table A1 – Results for P1. In this case, the *Spatials* conversion produces a plan that is superior to the plan produced with the *Current* method in all areas except the PTV maximum dose.

Structure	Metric	Original	Spatials Δ		Current Δ	
				Δ		Δ
PTV	$D_{98\%}[Gy]$	67.41	66.40	-1.01	65.77	-1.64
PTV	$D_{2\%}[Gy]$	70.83	71.29	0.46	71.35	0.52
PTV	$D_{max}[Gy]$	71.81	73.48	1.67	73.03	1.21
PTV	HI	0.05	0.07	0.02	0.08	0.03
Fem. Head, L	$D_{max}[Gy]$	49.50	47.42	-2.09	49.78	0.27
Fem. Head, R	$D_{max}[Gy]$	50.55	48.26	-2.29	56.26	5.71
Bladder	$V_{65Gy}[\%]$	27.33	27.46	0.13	27.36	0.03
Bladder	$V_{70Gy}[\%]$	2.41	4.10	1.69	5.38	2.97
Rectum	$V_{50Gy}[\%]$	26.39	27.63	1.24	31.24	4.85
Rectum	$V_{65Gy}[\%]$	9.71	9.30	-0.41	9.50	-0.22
Body	$D_{mean}[Gy]$	5.12	5.40	0.28	5.57	0.45
	HOT [cc]	28.02	23.45	-4.57	42.07	14.05
	RMSE [pp]		5.05	5.05	7.73	7.73

Table A2 – Results for P2. In this case the *Spatials* conversion results in significant increase in the homogeneity index due to a very low $D_{98\%}$. The values for bladder are very small ($< 2\%$) and thus not reported.

Structure	Metric	Original	Spatials Δ		Current Δ	
PTV	$D_{98\%}[Gy]$	81.29	77.05	-4.23	79.94	-1.34
PTV	$D_{2\%}[Gy]$	83.83	84.32	0.49	84.30	0.47
PTV	$D_{max}[Gy]$	84.17	85.17	0.99	85.28	1.11
PTV	HI	0.03	0.09	0.06	0.05	0.02
Fem. Head, L	$D_{max}[Gy]$	45.42	50.20	4.78	48.73	3.31
Fem. Head, R	$D_{max}[Gy]$	33.81	33.24	-0.57	38.61	4.80
Bladder	$V_{65Gy}[\%]$	1.91	1.74	-0.17	1.77	-0.15
Bladder	$V_{70Gy}[\%]$	1.54	1.37	-0.18	1.42	-0.12
Rectum	$V_{50Gy}[\%]$	8.28	10.18	1.90	10.16	1.88
Rectum	$V_{65Gy}[\%]$	4.37	4.73	0.36	5.16	0.78
Body	$D_{mean}[Gy]$	1.80	2.05	0.24	2.12	0.32
	HOT $[cc]$	36.76	28.14	-8.62	39.80	3.03
	RMSE $[pp]$		7.60	7.60	10.32	10.32

Table A3 – Results for P3. The plan produced with the *Spatials* method is superior to the plan produced with the *Current* method in all metrics except the bladder's V_{65Gy} , where both converted plans have lower volume than the original plan.

Structure	Metric	Original	Spatials Δ		Current Δ	
PTV	$D_{98\%}[Gy]$	66.61	65.39	-1.23	64.96	-1.66
PTV	$D_{2\%}[Gy]$	69.28	69.88	0.60	70.01	0.73
PTV	$D_{mean}[Gy]$	67.94	67.94	0.00	67.94	0.00
PTV	$D_{max}[Gy]$	70.64	72.39	1.75	71.83	1.20
PTV	HI	0.04	0.07	0.03	0.07	0.04
Fem. Head, L	$D_{max}[Gy]$	54.59	53.70	-0.89	54.71	0.12
Fem. Head, R	$D_{max}[Gy]$	48.37	45.56	-2.81	49.18	0.81
Bladder	$V_{65Gy}[\%]$	56.63	55.90	-0.73	54.01	-2.62
Bladder	$V_{70Gy}[\%]$	0.00	0.47	0.47	0.95	0.95
Rectum	$V_{50Gy}[\%]$	34.01	35.06	1.05	39.56	5.55
Rectum	$V_{65Gy}[\%]$	15.18	14.01	-1.18	13.95	-1.24
Body	$D_{max}[Gy]$	70.64	72.39	1.75	71.83	1.20
Body	$D_{mean}[Gy]$	5.52	5.96	0.43	6.35	0.83
	RMSE $[pp]$	0.00	5.16	5.16	8.58	8.58
	HOT $[cc]$	20.31	13.81	-6.50	30.22	9.92

B Plan Conversion tests: Head and Neck

Table B1 – Results for HN1. This is one of the two cases where the *Spatial*s method resulted in a clinically unacceptable hot spot inside the PTV High. The plan converted with the *Current* method results has smaller D_{max} for the PTV High. Interestingly, its body D_{max} is higher than the PTV High D_{max} , meaning that the highest dose goes outside the targets.

Structure	Metric	Original	Spatial Δ		Current Δ	
PTV High	$D_{98\%}[Gy]$	64.70	63.83	-0.87	63.35	-1.35
PTV High	$D_{2\%}[Gy]$	69.74	70.90	1.15	70.85	1.10
PTV High	$D_{mean}[Gy]$	67.52	67.52		67.52	
PTV High	$D_{max}[Gy]$	71.33	81.22	9.89	73.64	2.31
PTV High	HI	0.07	0.10	0.03	0.11	0.04
Brain	$D_{max}[Gy]$	1.02	1.19	0.17	1.18	0.16
Brainstem	$D_{2\%}[Gy]$	0.83	1.05	0.23	1.04	0.21
Brainstem	$D_{max}[Gy]$	1.01	1.22	0.21	1.21	0.20
Chiasm	$D_{max}[Gy]$	0.23	0.33	0.10	0.33	0.10
Ear, L	$D_{mean}[Gy]$	0.36	0.46	0.11	0.45	0.10
Ear, R	$D_{mean}[Gy]$	0.35	0.44	0.09	0.43	0.08
Esophagus	$D_{mean}[Gy]$	64.37	64.22	-0.14	64.94	0.57
Eye, L	$D_{2\%}[Gy]$	0.37	0.39	0.02	0.37	0.00
Eye, L	$D_{max}[Gy]$	0.38	0.54	0.16	0.52	0.14
Eye, R	$D_{2\%}[Gy]$	0.37	0.36	-0.01	0.35	-0.02
Eye, R	$D_{max}[Gy]$	0.38	0.50	0.12	0.47	0.09
Larynx	$D_{mean}[Gy]$	67.56	67.85	0.29	67.55	-0.00
Mandible	$D_{2\%}[Gy]$	47.71	48.99	1.28	47.35	-0.36
Mandible	$D_{max}[Gy]$	53.72	56.67	2.95	54.87	1.15
O. Nerve, L	$D_{max}[Gy]$	0.38	0.26	-0.11	0.26	-0.12
O. Nerve, R	$D_{max}[Gy]$	0.37	0.30	-0.07	0.30	-0.07
Parotid, L	$D_{mean}[Gy]$	2.35	2.53	0.18	2.59	0.24
Parotid, R	$D_{mean}[Gy]$	1.42	1.56	0.13	1.60	0.18
Spinal Cord	$D_{2\%}[Gy]$	36.60	35.86	-0.74	37.06	0.47
Spinal Cord	$D_{max}[Gy]$	38.50	39.30	0.80	43.36	4.86
Body	$D_{max}[Gy]$	71.33	81.22	9.89	77.10	5.77
Body	$D_{mean}[Gy]$	8.36	7.96	-0.40	8.09	-0.26
	RMSE [pp]		4.83	4.83	7.59	7.59

Table B2 – Results for HN2. In this case the *Spatials* conversion resulted in better values for almost all metrics when compared to the *Current* conversion. The only clinically relevant exception is the D_{max} for the PTV High. The RMSE values show that of the two plans, the dose distribution of the plan created with the *Spatials* conversion is significantly closer to the original dose distribution.

Structure	Metric	Original	Spatials Δ		Current Δ	
PTV High	$D_{98\%}[Gy]$	58.28	57.84	-0.43	56.96	-1.32
PTV High	$D_{2\%}[Gy]$	63.45	64.43	0.97	64.40	0.94
PTV High	$D_{mean}[Gy]$	61.55	61.55		61.55	
PTV High	$D_{max}[Gy]$	64.60	68.26	3.67	67.23	2.63
PTV High	HI	0.08	0.11	0.02	0.12	0.04
PTV Med	$D_{98\%}[Gy]$	50.38	49.93	-0.45	50.03	-0.35
PTV Med	$D_{2\%}[Gy]$	55.54	56.45	0.91	56.70	1.16
PTV Med	$D_{max}[Gy]$	53.32	53.35	0.03	54.06	0.74
PTV Med	$D_{mean}[Gy]$	57.42	60.34	2.93	61.41	3.99
PTV Med	HI	0.10	0.12	0.03	0.12	0.03
Brain	$D_{2\%}[Gy]$	26.19	25.79	-0.39	28.86	2.68
Brain	$D_{max}[Gy]$	48.43	51.39	2.95	52.73	4.29
Brainstem	$D_{2\%}[Gy]$	33.64	33.41	-0.23	35.00	1.36
Brainstem	$D_{max}[Gy]$	39.23	37.32	-1.91	44.25	5.02
Chiasm	$D_{2\%}[Gy]$	2.34	2.36	0.02	2.39	0.05
Chiasm	$D_{max}[Gy]$	2.42	2.44	0.02	2.47	0.05
Ear, L	$D_{mean}[Gy]$	14.20	16.16	1.95	13.95	-0.25
Ear, R	$D_{mean}[Gy]$	8.66	9.77	1.11	8.47	-0.19
Esophagus	$D_{mean}[Gy]$	24.88	25.13	0.25	25.02	0.14
Eye, L	$D_{2\%}[Gy]$	1.56	1.58	0.02	1.66	0.10
Eye, L	$D_{max}[Gy]$	1.76	1.78	0.01	1.87	0.10
Eye, R	$D_{2\%}[Gy]$	1.59	1.62	0.02	1.53	-0.06
Eye, R	$D_{max}[Gy]$	1.73	1.76	0.03	1.66	-0.07
Larynx	$D_{mean}[Gy]$	23.91	24.62	0.71	25.62	1.71
Mandible	$D_{2\%}[Gy]$	58.71	58.75	0.03	62.24	3.53
Mandible	$D_{max}[Gy]$	63.57	63.97	0.41	64.70	1.13
O. Nerve, L	$D_{2\%}[Gy]$	1.92	1.94	0.02	2.00	0.08
O. Nerve, L	$D_{max}[Gy]$	2.01	2.03	0.01	2.08	0.07
O. Nerve, R	$D_{2\%}[Gy]$	2.11	2.14	0.03	2.15	0.04
O. Nerve, R	$D_{max}[Gy]$	2.35	2.39	0.04	2.43	0.08
Parotid, L	$D_{mean}[Gy]$	39.28	40.12	0.84	39.82	0.55
Parotid, R	$D_{mean}[Gy]$	8.16	8.91	0.74	10.16	2.00
Spinal Cord	$D_{2\%}[Gy]$	34.54	33.05	-1.49	38.91	4.37
Spinal Cord	$D_{max}[Gy]$	38.13	36.21	-1.92	46.67	8.54
Body	$D_{max}[Gy]$	64.60	68.26	3.67	67.23	2.63
Body	$D_{mean}[Gy]$	7.87	7.89	0.03	8.49	0.62
Body	HOT [cc]	2.80	4.41	1.61	17.81	15.01
	RMSE [pp]		2.19	2.19	8.09	8.09

Table B3 – Results for HN3. In this case the *Spatial*s conversion resulted in an significant increase in the D_{max} for mandible. The goal for the mandible’s D_{max} is < 70 Gy and the increase is significant. However, the *Current* conversion method results also in a significant increase for the same metric. For $D_{2\%}$ the increase is much smaller.

Structure	Metric	Original	Spatial Δ		Current Δ	
PTV High	$D_{98\%}[Gy]$	69.90	69.64	-0.26	69.36	-0.54
PTV High	$D_{2\%}[Gy]$	75.57	75.85	0.28	76.18	0.62
PTV High	$D_{mean}[Gy]$	72.74	72.74	0.00	72.74	0.00
PTV High	$D_{max}[Gy]$	77.40	81.01	3.61	79.05	1.65
PTV High	HI	0.08	0.09	0.01	0.09	0.02
PTV Med	$D_{98\%}[Gy]$	64.85	64.16	-0.69	64.40	-0.44
PTV Med	$D_{2\%}[Gy]$	71.61	71.55	-0.06	73.28	1.68
PTV Med	$D_{max}[Gy]$	68.44	68.18	-0.26	69.29	0.85
PTV Med	$D_{mean}[Gy]$	74.45	74.59	0.13	76.69	2.24
PTV Med	HI	0.10	0.11	0.01	0.13	0.03
Brain	$D_{2\%}[Gy]$	42.04	41.69	-0.35	48.14	6.10
Brain	$D_{max}[Gy]$	73.79	73.55	-0.24	72.22	-1.57
Brainstem	$D_{2\%}[Gy]$	16.18	17.73	1.54	17.43	1.25
Brainstem	$D_{max}[Gy]$	24.07	24.61	0.55	24.24	0.18
Chiasm	$D_{max}[Gy]$	2.40	3.01	0.61	3.39	1.00
Ear, L	$D_{mean}[Gy]$	29.80	31.61	1.81	26.98	-2.82
Ear, R	$D_{mean}[Gy]$	6.98	8.48	1.50	7.01	0.03
Esophagus	$D_{mean}[Gy]$	35.93	36.90	0.98	34.18	-1.74
Eye, L	$D_{2\%}[Gy]$	2.33	2.98	0.65	3.08	0.74
Eye, L	$D_{max}[Gy]$	2.53	3.16	0.63	3.37	0.84
Eye, R	$D_{2\%}[Gy]$	1.67	2.43	0.76	2.03	0.36
Eye, R	$D_{max}[Gy]$	1.90	2.66	0.75	2.19	0.29
Larynx	$D_{mean}[Gy]$	29.13	30.23	1.11	29.27	0.14
Mandible	$D_{2\%}[Gy]$	72.71	73.71	1.00	75.06	2.35
Mandible	$D_{max}[Gy]$	74.98	81.01	6.03	79.05	4.07
O. Nerve, L	$D_{max}[Gy]$	3.14	3.57	0.43	3.79	0.65
O. Nerve, R	$D_{max}[Gy]$	2.63	3.27	0.64	3.26	0.63
Parotid, R	$D_{mean}[Gy]$	2.32	3.46	1.14	4.94	2.62
Spinal Cord	$D_{2\%}[Gy]$	36.25	35.54	-0.71	38.29	2.04
Spinal Cord	$D_{max}[Gy]$	38.47	39.41	0.94	43.83	5.36
Body	$D_{max}[Gy]$	77.40	81.01	3.61	79.05	1.65
Body	$D_{mean}[Gy]$	10.16	10.37	0.21	10.96	0.80
Body	HOT [cc]	76.20	69.84	-6.36	125.36	49.15
	RMSE [pp]		3.98		9.94	